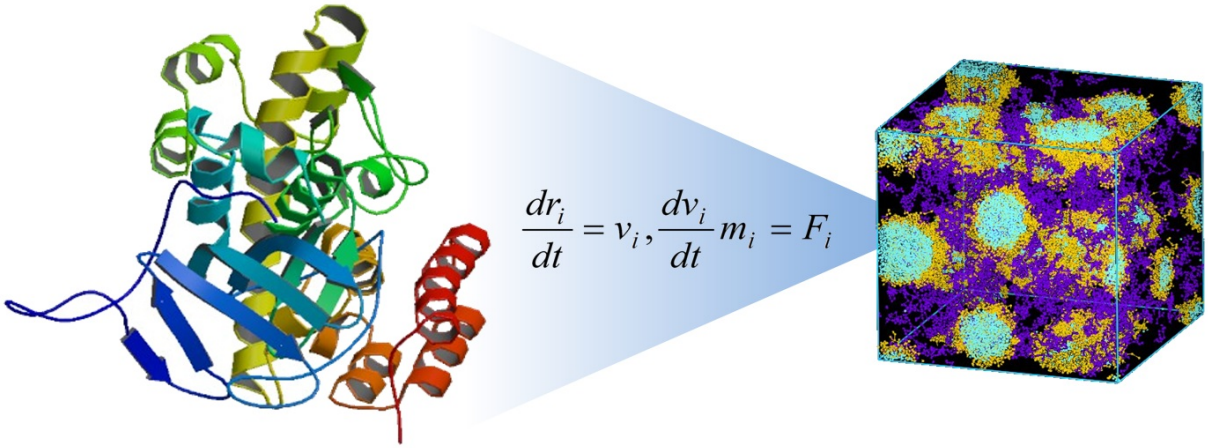
Simulation von Peptiden, Proteinen und Biomembranen mit Molekularer Fragmentdynamik (MFD)



Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

– Dr. rer. nat. –

vorgelegt von

Andreas Truszkowski

geboren in Recklinghausen

Fakultät für Chemie der

Universität Duisburg-Essen

13. November 2014

Für meine Familie...

Die vorliegende Arbeit wurde im Zeitraum von 09.2011 bis 11.2014 im Arbeitskreis von Prof. Dr. Matthias Epple am Institut für Anorganische Chemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 30.01.2015

Gutachter:	Herr Prof. Dr. rer. nat. Matthias Epple
	Herr Prof. Dr. rer. nat. Achim Zielesny
Vorsitzender:	Herr Prof. Dr. rer. nat. Eckhard Spohr

Inhaltsverzeichnis

1 Einleitung

Seit der Entwicklung des Computers gehören computergestützte Simulationen zu den wichtigen Werkzeugen in der Naturwissenschaft. Sie helfen komplexe Vielteilchenprobleme[1] zu bearbeiten, um so z.B. Theorien wie die Newtonsche Mechanik auf molekulare Systeme anzuwenden. Computersimulationen können dabei die Lücke zwischen Theorie und Experiment schließen. Die ersten computergestützten Untersuchungen von Molekülen und einfachen chemischen Systemen fanden in den 1950er und 1960er Jahren statt[2–5]. Seither hat sich die Rechenleistung um einen Faktor von 10^{10} gesteigert[6,7], sodass immer größere molekulare Ensembles mit genaueren Modellen untersucht werden können[8–11]. Computersimulationen bekommen daher einen immer größeren Stellenwert in der Chemie (Nobelpreise 1998[12] und 2013[13]) und insbesondere der molekularen Biologie. Die gängigsten atomistischen Methoden für die strukturelle Untersuchung von Biomolekülen sind Monte-Carlo-(MC)-[14–17] und Molekulardynamik-(MD)-Verfahren[18–22] sowie eine Mischung aus Kraftfeldmethoden und quantenmechanischen *ab-initio*-Verfahren (Quantenmechanik/Molekularmechanik, kurz QM/MM)[23–26].

Ein grundlegendes Problem der genannten computergestützten Methoden sind die großen Längen- und Zeitskalen für die Untersuchung von molekularbiologischen Systemen. Atomistische Simulationen sind nur begrenzt einsetzbar, um molekulare Strukturen mit Systemgrößen von Mikrometern und Simulationszeiten von Mikrosekunden in einer angemessenen Rechenzeit zu analysieren.

Ein möglicher Ansatz, um in die Größenordnungen mesoskaler molekularbiologischer Systeme vorzudringen, sind sogenannte *coarse-grained* Verfahren (deutsch etwa „grobe Körnung“, kurz CG). Ein CG-Modell mittelt die Details des atomis-

tischen Systems in größere Einheiten. Hierdurch kommt es zu einer Reduktion der vorhandenen Freiheitsgrade und der Eliminierung von atomistischen Wechselwirkungen, was bei entsprechenden Simulationen zu einer Geschwindigkeitssteigerung von mehreren Größenordnungen führt[18,27–31]. Ein weiteres Problem der atomistischen Verfahren ist ihre schwierige praktische Handhabung. Aufgrund der harten atomaren Potentiale, welche das wechselseitige Durchdringen molekularer Strukturen verhindern, führen ungünstige Startgeometrien häufig zu Verdrillungs- bzw. Käfigeffekten von Molekülen innerhalb der Simulationen. Dies hat zur Folge, dass die Relaxation und Energieminimierung durch sterische Behinderungen der Moleküle untereinander gestört, und damit verlangsamt wird[32]. Um dieses Phänomen zu umgehen, sind zeitaufwendige Design- und Optimierungsschritte der betrachteten Modelle im Vorfeld der eigentlichen Analyse vonnöten.

Die dissipative Partikeldynamik (DPD) ist eine mesoskopische *coarse-grained* Simulationstechnik zur Analyse des hydrodynamischen Verhaltens von isothermalen Flüssigkeiten in großen Längen- und Zeitskalen, wobei chemische Ensembles mit mehreren Millionen Atomen betrachtet werden können[33]. Eingeführt wurde die DPD-Methode in den frühen 1990ern von Hoogerbrugge und Koelman[34,35]. Die grundlegenden wechselwirkenden Teilchen, in der DPD-Theorie auch *Beads* genannt, entsprechen nicht mehr den Elektronen oder Atomkernen der quantenmechanischen Verfahren bzw. den feinkörnigen (engl. *fine-grained*) Atomtypen der MM/MD-Methoden, sondern größeren molekularen Einheiten. Die *Beads* unterliegen weichen Potentialen, sodass eine wechselseitige Durchdringung prinzipiell möglich ist. Dies führt zu einem Geschwindigkeitsvorteil gegenüber Techniken mit harten Teilchenpotentialen, da es zu einer schnelleren Relaxation des Simulationssystems kommt. Außerdem ist die Startgeometrie weniger anfällig gegenüber geometrischen Verschachtelungen der einzelnen Moleküle, was das Design von Simulationen vereinfacht[33,36–40].

Motivation und Programm der vorliegenden Arbeit

In der dissipativen Partikeldynamik sind die *Beads* nicht auf einzelne Moleküle beschränkt, sondern können jede Region in einer Flüssigkeit darstellen. Zusätzlich bildet ein *Bead* häufig nicht nur eine diskrete molekulare Einheit ab, sondern repräsentiert viele Einheiten mit vergleichbarem chemischen Charakter[41]. Die hier verwendete Variante der dissipativen Partikeldynamik - die molekulare Fragmentdynamik (MFD) - verfolgt den Ansatz, dass jedes *Bead* grundsätzlich ein individuelles Molekül oder Molekülfragment widerspiegelt. Jedes Molekül in einer MFD-Simulation wird aus einem oder mehreren Molekülfragmenten, welche isomorph zu einem Teil oder der gesamten molekularen Struktur sind, aufgebaut. Die Bindungen zwischen den einzelnen Fragmenten werden anhand von harmonischen Potentialen abgebildet. Die Gesamtheit aller Molekülfragmente einer MFD-Simulation wird als Fragmentset bezeichnet. Das in dieser Arbeit entworfene Fragmentset wird für die Darstellung von Proteinen, Peptiden und Biomembranen konzeptioniert. Trotz der gezielten Entwicklung ist das Fragmentset weit gefächert anwendbar und nicht nur auf molekularbiologische Systeme beschränkt. Für die interne Repräsentation von molekularen Strukturen aus Fragmenten wird eine einzeilige Schreibweise zur Charakterisierung der Fragmenttopologie (auch Fragmentierungsschema genannt) auf Basis der bewährten SMILES-Notation[42–44] eingeführt und als *fragment* SMILES oder kurz als *f*SMILES bezeichnet. Dieser Ansatz ermöglicht den Aufbau großer makromolekularer Strukturen aus einfachen chemischen Einheiten, den Molekülfragmenten. Die zwingende fragmentbasierte Repräsentation der Simulationen macht die molekulare Fragmentdynamik zu einer chemisch intuitiven Methode.

Für die fragmentbasierte Beschreibung von Proteinen wird ein Fragmentierungsschema für alle 20 proteinogenen Aminosäuren erarbeitet. Die manuelle Generierung eines *f*SMILES ist für niedermolekulare Strukturen, die nur aus wenigen Molekülfragmenten zusammengesetzt sind, nicht aufwendig. Hingegen beträgt die durchschnittliche Länge der Aminosäuresequenz eines Proteins 320 Aminosäuren[45]. Diese bestehen im hier entwickelten MFD-Fragmentset im Mittel aus drei Fragmenten. Zusätzlich soll der Gebrauch von Protein-Data-Bank-(PDB)-

Archivdateien[46] den Einsatz einer von der Kristallstruktur abgeleiteten dreidimensionalen Proteinstruktur in MFD-Simulationen ermöglichen. Die Bereitstellung eines *f*SMILES und die Verknüpfung mit Positionsdaten der einzelnen Fragmente ist aufgrund der Datenmenge per Hand in angemessener Zeit nicht möglich. Hierfür wird im Rahmen dieser Arbeit eine Softwarebibliothek erarbeitet, welche die Verwaltung der Daten übernimmt und für die Simulation aufbereitet. Zur Darstellung der Proteindaten sowie der fragmentbasierten Simulationsboxen wird ein Softwaremodul auf Basis von Jmol[47] entwickelt, einem *Open-Source*-Projekt zur interaktiven Visualisierung von molekularen Strukturen.

Für jedes Molekülpaar eines Fragmentsets ist der DPD-Repulsionsparameter der konservativen Kraft zu bestimmen, welcher alle physikalischen nicht-bindenden Wechselwirkungen eines Fragmentpaares gemittelt beschreibt. In vielen DPD-Modellen wird der Repulsionsparameter auf ein spezifisches Problem hin angepasst. Dies gewährleistet, das zu erzielende Verhalten der *Beads* zueinander zu erreichen, aber ignoriert detaillierte molekulare Wechselwirkungen wie die Van-der-Waals-Kräfte[41]. Für das hier aufgestellte Fragmentset sollen die Repulsionsparameter weitläufig anwendbar sein. Aus diesem Grund werden die zu bestimmenden Repulsionsparameter aus atomistischen MD-Simulationen generiert[48–53]. Für die so entstehenden Datenmengen (Größenordnung: 10^5 Energiewerte) zur Berechnung der Parameter wird eine spezielle Verwaltungssoftware bereitgestellt. Zusätzlich soll die Software die Eigenschaften der Molekülfragmente sowie Topologiedaten der Aminosäuren zur Beschreibung von Proteinen strukturiert organisieren. Das so erstellte Fragmentset wird auf alle in dieser Arbeit behandelten biomolekularen Probleme angewendet.

In dieser Arbeit wird die noch in der Entwicklung befindliche Simulationssoftware MFD-FormulaOne[54] angewendet. Zum Test der Software wird eine initiale MFD-Studie von Polyoxyethylenalkylether-Tensiden C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$ durchgeführt. Die Studie soll zeigen, dass die in der experimentellen Literatur beschriebenen nanoskaligen Strukturen im thermodynamischen Gleichgewicht dargestellt werden können. Die überwiegende Anzahl an Studien betrachten Tensidmischungen an einer Wasser-Öl-Grenzfläche[55–66]. In der hier

vorgelegten Abhandlung wird eine Wasser-Luft-Grenzfläche eingeführt und das Verhalten der Tenside an dieser Grenzfläche analysiert. Zusätzlich werden die errechneten Werte für die Oberflächenspannung sowie der kritischen Mizellbildungskonzentration (CMC) der betrachteten Tenside mit experimentellen Werten verglichen.

Für eine genäherte Beschreibung der elektrostatischen Wechselwirkung innerhalb von MFD-Simulationen von Peptiden, Proteinen und Biomembranen wird ein einfaches elektrostatisches Ad-Hoc-Konzept entwickelt. Die molekulare Fragmentdynamik berücksichtigt grundsätzlich nicht langreichweitige Wechselwirkungen. Damit die harten Potentiale der elektrostatischen Wechselwirkung vereinbar mit den weichen MFD-Potentialen sind, ist eine gedämpfte Implementierung notwendig. Bei der Anwendung ist darauf zu achten, dass es zu einem natürlichen Lösungsverhalten der Ionen zwischen freien und aggregierten Teilchen kommt. Die Bildung von künstlichen Ionenclustern ist zu vermeiden[67].

Die nicht-bindenden chemischen Wechselwirkungen werden in MFD-Simulationen mit Hilfe des isotropen DPD-Repulsionsparameters angenähert. Proteine hingegen erhalten ihre Struktur durch anisotrope gerichtete Wechselwirkungen wie Wasserstoffbrückenbindungen. Um dennoch die Proteinkonformation aufrecht erhalten zu können, werden harmonische Stabilisierungspotentiale zwischen den α-Kohlenstofffragmenten des Proteinrückgrates einfügt und untersucht. Die so erhaltene Stabilisierung des Proteinrückgrates soll in ihrer Stärke flexibel bis hin zu starr frei einstellbar sein. Als Modellsystem wird das hantelförmige Protein Calmodulin untersucht. Hierbei wird die genaue Auswirkung der Stabilisierungspotentiale auf konformelle Änderungen des Proteins betrachtet. Damit die geometrischen Veränderungen nachzuvollziehen und auszuwerten sind, werden sogenannte Sondenfragmente in das Proteinrückgrat eingebettet.

Zinkricinoleat, $Zn(Ri)_2$, ist aufgrund seiner minimalen Toxizität ($LD_{50} < 2000$ $mg{\cdot}kg^{-1}$) sowie seiner geringen sensitivierenden Wirkung auf Haut und Augen interessant für die Anwendung in Haushalt, Kosmetik und pharmazeutischen Industrie[68–72]. Für Zinkricinoleat ist bekannt, dass es als Medium zur Aufreinigung von polyhistidinhaltigen (His-Tag-) Proteinen verwendbar ist[69,73,74]. In

der vorliegenden Arbeit wird die Modellsubstanz Zinkricinoleat auf die gezielte Bindungsfähigkeit von schwefel- und stickstoffhaltigen funktionellen Gruppen hin untersucht. Für die Untersuchung der oberflächenaktiven Eigenschaften des Zinkricinoleates muss zuerst eine stabile $Zn(Ri)_2$-Schicht in einer MFD-Simulation abgebildet werden. Anschließend wird die Wechselwirkung von Aminosäuren (Histidin versus alternative Aminosäure) und Polyhistidinpeptiden mit den modellierten Schichten betrachtet. Zur Darstellung der spezifischen Bindungsfähigkeit des Zinkricinoleates zum stickstoffhaltigen Molekülfragment Imidazol wird in einer kompetitiven Simulation gleichzeitig das Verhalten von Proteinen mit und ohne His-Tag untersucht.

Als weiteres molekulares Schichtmodellsystem sollen Phospholipiddoppelschichten mit der molekularen Fragmentdynamik simuliert werden. Diese nehmen eine bedeutende Stellung in biologischen Systemen ein, um gesonderte Kompartimente bzw. Reaktionsräume innerhalb oder zwischen einzelnen Zellen zu unterteilen[75–78]. Die Mehrheit an computergestützten Simulationen untersuchen einfache Membranmodelle mit lediglich einer Spezies von Phospholipiden auf deren Stabilität, Zusammensetzung, Struktur und Fusion[79–82]. Die Simulation einer einfachen 1,2-Dimyristoyl-sn-glycero-3-phosphocholin-(DMPC)-Membran soll zeigen, dass es möglich ist eine stabile Doppelschicht ohne Ausbildung von Mizellen und Löchern zu realisieren. Charakteristische Membrankenngrößen wie hydrophobe Dicke, Gesamtdicke und Fläche pro Lipid werden mit Literaturangaben verglichen. Nach der Beschreibung der DMPC-Membran sollen die Eigenschaften von komplexeren Membranen, nämlich der des endoplasmatischen Retikulums (ER), des Mitochondriums und Plasmamembranen, ausgewertet werden. Auch hier werden die Membrancharakteristika mit der Literatur verglichen[76,83–85]. Die spontane Bildung von Mizellen, Vesikeln und Liposomen ist eine weitere allgemeine Eigenschaft von Phospholipiden[86,87]. Dieses Verhalten sowie die Selbstorganisation in die genannten Strukturen wird im Anschluss sowohl für das einfache DMPC-Modell als auch für die drei komplexeren Membrantypen erforscht. Ein anderer wichtiger Prozess in lebenden Zellen ist die Verschmelzung von Phospholipiddoppelschichten. Dieser Vorgang ist grundlegend für Mechanismen wie dem zellularen Transport mittels Vesikeln oder der Signalübertragung

an den Synapsen[88]. In einer Simulation soll der Fusionsprozess eines Vesikels mit einer DMPC-Membran dargelegt werden.

Nach der Einführung von biologischen Membranen in der molekularen Fragmentdynamik wird die Interaktion dieser Lipiddoppelschichten mit einem Protein genauer betrachtet. Für das Protein Kalata B1 - ein Pflanzendefensin aus der Gruppe der Cyclotide - soll der Mechanismus der zytotoxischen Eigenschaften charakterisiert werden. Betrachtet wird in mehreren MFD-Simulationen eine Plasmamembran, welche unterschiedlichen Konzentration (24 - 350 Proteine) des Proteins Kalata B1 ausgesetzt wird. Im ersten Schritt wird der in der Literatur vorgeschlagene Ablauf des Membrankontaktes über spezifische Proteindomänen studiert[89–94]. Im Anschluss soll die Zerstörung der Membranstruktur anhand von Krümmungsstress und der zusätzlich auftretenden Lipidextraktion in Abhängigkeit der Kalata B1 Konzentration bestimmt werden.[95,96]. In der Literatur werden zwei mögliche Arten von Porenstrukturen beschrieben: Die Fasslamellenpore und die Ringkernpore[89,92,95]. Die Art und Struktur der Pore sollen ermittelt werden.

2 Theoretische Grundlagen

Dieses Kapitel soll eine kurze Einführung der verwendeten Methoden skizzieren und dem Leser zu verdeutlichen, auf welcher theoretischen Basis die hier vorliegende Arbeit aufbaut. Der erste Teil erläutert alle verwendeten computergestützten Simulationsverfahren. Im zweiten Abschnitt werden die chemischen Modellsysteme genauer beschrieben.

2.1 Molekulare Modellierung

Die molekulare Modellierung (engl. *Molecular Modelling*) fasst computergestützte Techniken zusammen, die versuchen das Verhalten von molekularen Systemen durch mathematische Modelle zu beschreiben[97]. Aufgrund des hohen Rechenaufwandes für die Beschreibung chemischer Modelle profitiert die molekulare Modellierung heutzutage von den Fortschritten im Bereich der Computersoftware und Hardware[98]. Die ersten computergestützten Methoden die untersucht wurden, waren Monte-Carlo-(MC)-[5] und Molekulardynamik-(MD)-Simulationen[2–4] in den 1950er und 1960er Jahren[33]. Die Rechengeschwindigkeit wurde seitdem um den Faktor 10^{10} gesteigert, wie es das Mooresche Gesetz[6] aus dem Jahre 1965 vorhersagte. Hieraus ergibt sich, dass immer genauere theoretische Betrachtungen der molekularen Systeme möglich sind sowie immer größere chemische Systeme untersucht werden können[8–11].

Ungeachtet der Steigerung der Rechenleistung in den letzten Jahrzehnten ist es dennoch nötig, die Komplexität der Berechnungen zu vereinfachen, wenn mesoskale Systeme in angemessener Zeit betrachtet werden sollen. Die aufwendigsten

und zugleich auch realistischsten Simulationen sind Berechnungen auf Basis von quantenmechanischen *ab-initio*-Verfahren. Quantenmechanische Verfahren basieren auf der Lösung der Schrödingergleichung. Dies ist für größere Systeme nur numerisch und unter der Anwendung von Approximationsverfahren wie der Born-Oppenheimer-Näherung möglich. Auch unter Verwendung dieser Vereinfachungen sind quantenmechanische Verfahren nur auf kleinste Systeme (Größenordnung: 10^{-10} m) und im Bereich von Pikosekunden einsetzbar. Eine der aufwendigsten *ab-initio*-Berechnungen, die bis heute auf MP2-Niveau durchgeführt wurden, ist die Berechnung der Eigenschaften einer kubischen Simulationsbox mit insgesamt 64 Wassermolekülen[99]. Diese Größenordnung ist weit entfernt von der Simulation von molekularbiologischen Systemen wie Biomembranen, Peptiden, Proteinen oder der Kombinationen der einzelnen Systeme.

Für chemische Systeme im Nanometerbereich sind prinzipiell moderne Molekularmechanik-(MM)-Methoden anzuwenden. Die Molekülbewegung wird von Newtons zweitem Axiom bestimmt. Die kleinste betrachtete Einheit ist das Atom. Harmonische Potentiale ersetzen die chemische Bindung. Die potentielle Energie des Systems wird aus den Parametern eines sogenannten Kraftfeldes errechnet, welche mittels theoretisch komplexeren Rechnungen oder aus experimentellen Daten berechnet werden[100]. Die Zeitskalen reichen von Nano- bis zu Mikrosekunden und Längenmaßen von Nanometern. Biologische Systeme sind ein stetig wachsender Anwendungsbereich der MM-Methoden. Die Rechenzeiten erstrecken sich jedoch über Wochen bis hin zu Monaten[19,21].

Die im Rahmen dieser Arbeit angewendete Methode - die molekulare Fragmentdynamik (MFD) - ist eine mesoskopische Simulationstechnik auf Basis der dissipativen Partikeldynamik (DPD). Die molekulare Fragmentdynamik beruht wie die Kraftfeldmethoden auf dem zweiten Newtonschen Axiom zur Berechnung der Bewegung der involvierten Teilchen (*Beads*). Die kleinste Einheit sind in der molekularen Fragmentdynamik nicht Atome wie bei den vorher genannten MM-Methoden, sondern explizit Molekülfragmente oder ganze Moleküle. Verfahren mit einer reduzierten Anzahl an Freiheitsgraden und einem geringeren Detailgrad des betrachteten molekularen Systems nennt man *coarse-grained* Simula-

tionen. Die Technik des *coarse-grained* Ansatzes erhöht die Rechengeschwindigkeit um Größenordnungen gegenüber den atomistischen Verfahren (Stunden bis Tage). MM-Methoden benötigen häufig auf Grund des hohen Rechenaufwandes Supercomputer. Alle Rechnungen dieser Arbeit werden dagegen auf einer Intel Xeon™ Workstation durchgeführt. Die Simulationszeit reicht von Mikro- bis hin zu Millisekunden. Die Größenskalen dieser Technik reichen bis in den Mikrometerraum[33,48].

Abbildung 2.1 zeigt die Einteilung der verschiedenen Methoden der molekularen Modellierung in Abhängigkeit der jeweiligen Zeit- und Längenskalen.

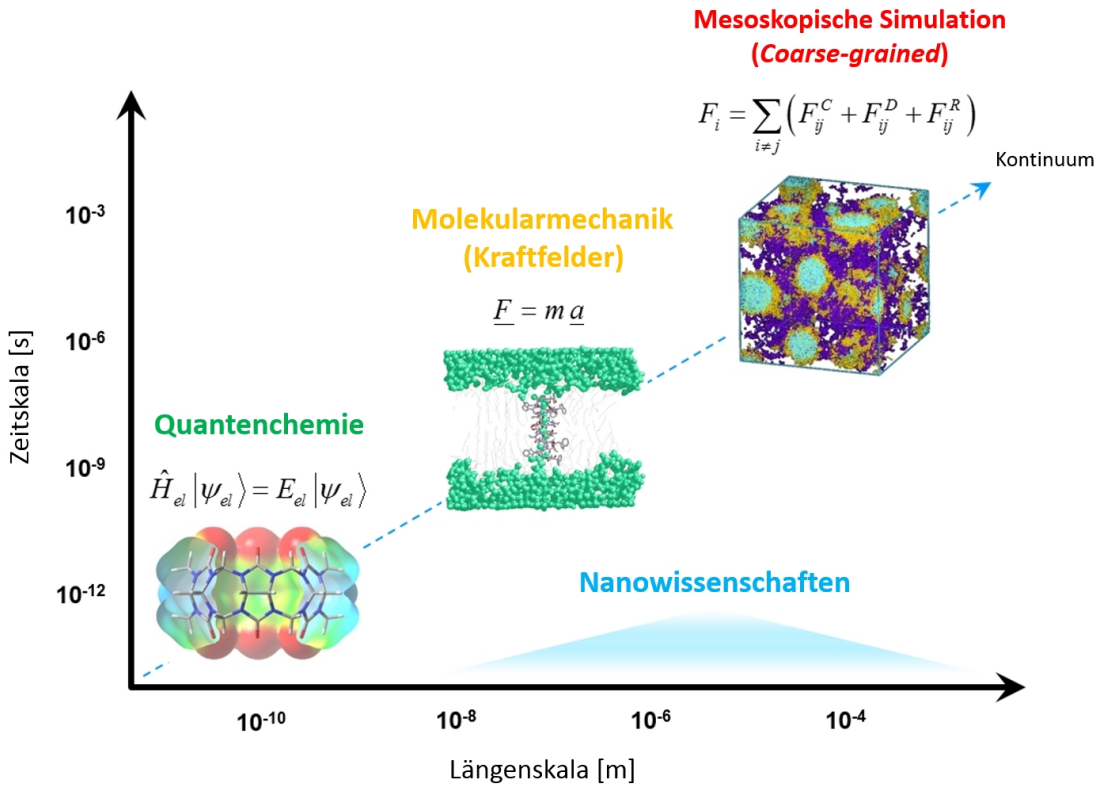


Abbildung 2.1: *Einteilung der verschiedenen computergestützten Simulationsverfahren anhand der jeweiligen Zeit- und Längenskalen.*

2.1.1 Molekülmechanik (MM)

Eines der Hauptprobleme der Quantenmechanik ist die Berechnung der Potentialhyperfläche für eine gegebene atomare Konfiguration. Molekülmechanik-

(MM)-Methoden, häufig auch als Kraftfeldmethoden bezeichnet, umgehen dieses Problem, indem sie das molekulare System als parametrisierte Funktionen der Raumkoordinaten der Atome abbilden. Die Parameter der Kraftfeldfunktionen werden dabei an experimentelle Daten oder an physikalisch exaktere Computersimulationen angepasst. Moleküle werden mittels eines Ball-und-Feder-Models beschrieben, wobei die Atome in unterschiedliche Atomtypen eingeteilt werden. Die Kraftfeldmethoden arbeiten auf der atomaren Ebene und werden als atomistische Verfahren bezeichnet. Die Details der Elektron-Elektron- und der Elektron-Nukleon-Wechselwirkungen werden vernachlässigt[100]. Die chemische Bindung, welche den Zusammenhalt der Atome vermittelt, wird als harmonische Feder (Hookesches Gesetz) mit variierender Länge charakterisiert[101].

Die Gesamtenergie E_{KF} eines Kraftfeldes kann als Summe der einzelnen Energieterme wiedergegeben werden:

$$E_{KF} = \sum E_l + \sum E_w + \sum E_t + \sum E_{vdw} + \sum E_{el} + \sum E_k \tag{2.1}$$

E_l ist die Bindungslängenenergie, E_w beschreibt die notwendige Bindunswinkelenergie, um einen Bindungswinkel zu ändern. E_t stellt die erforderliche Torsionswinkelenergie für eine Drehung um eine Bindung dar. E_{vdw} und E_{el} sind die nichtbindenden Atom-Atom-Wechselwirkungen und E_k bezeichnet man als Kopplungsterme (engl. *cross-terms*), welche Kopplungen zwischen den ersten drei Valenztermen beschreiben[101]. Abweichungen zu den im Kraftfeld hinterlegten Parametern erhöhen die Energie des Gesamtsystems[102]. In Kapitel 2.1.3 ist eine genaue Aufstellung der Energieterme für das COMPASS-Kraftfeld[100] aufgeführt, welches in dieser Arbeit zur Beschreibung der molekularen Systeme verwendet wird.

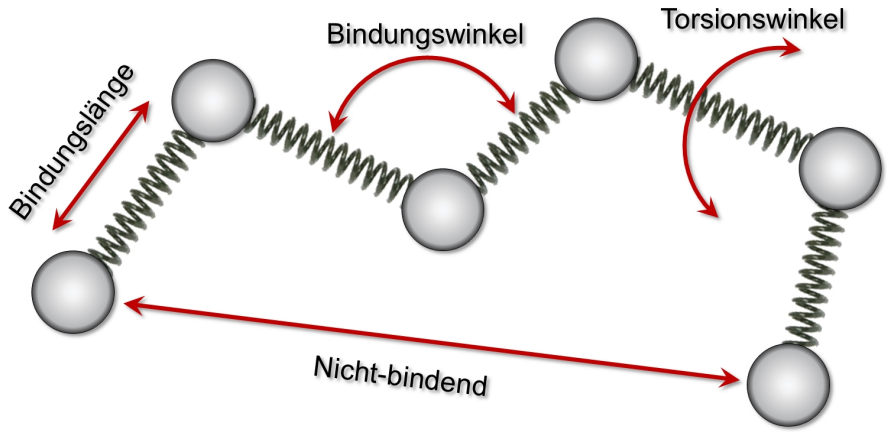


Abbildung 2.2: *Skizzierung der grundlegenden Energieterme eines Kraftfeldes*[101].

Die Raumkoordinaten der Atome des globalen Minimums der Potentialfunktionen des Kraftfeldes beschreiben das geometrische Energieminimum des Moleküls bzw. des molekularen Systems. Numerische Berechnungsmethoden für das Minimum des betrachteten chemischen Ensembles sind die Methoden *Steepest-Descent* (SD) und *Conjugate-Gradient* (CG). Häufig wird für eine grobe Annäherung an das Minimum das SD-Verfahren verwendet und die letzten Minimierungsschritte mittels der CG-Methode abgeschlossen.

Steepest-Descent (SD)

Das *Steepest-Descent*-(SD)-Verfahren ist ein Minimierungsverfahren, welches für Optimierungsprobleme ohne Nebenbedingungen einsetzbar ist:

$$\min \; f(x), \quad mit \; x \in R^n \tag{2.2}$$

wobei R^n ein n-dimensionaler euklidischer Raum ist und $f : R^n \to R$ eine kontinuierlich differenzierbare Funktion darstellt. Die Suchrichtung wird derart gewählt, sodass sie streng bergab des Gradienten $-\nabla f(x_k)$ verläuft. In Abbildung 2.3 ist der Pfad einer Minimierung skizziert. Jede Iteration produziert eine neue Suchrichtung, welche orthogonal zur vorherigen verläuft. Dabei oszilliert der Pfad auf dem Weg zum Minimum und schießt kontinuierlich über den besten Pfad hinaus,

woraus ein ineffizienter Bewegungsablauf entsteht. Dadurch konvergiert das SD-Verfahren nur langsam in der Nähe des Minimums[103].

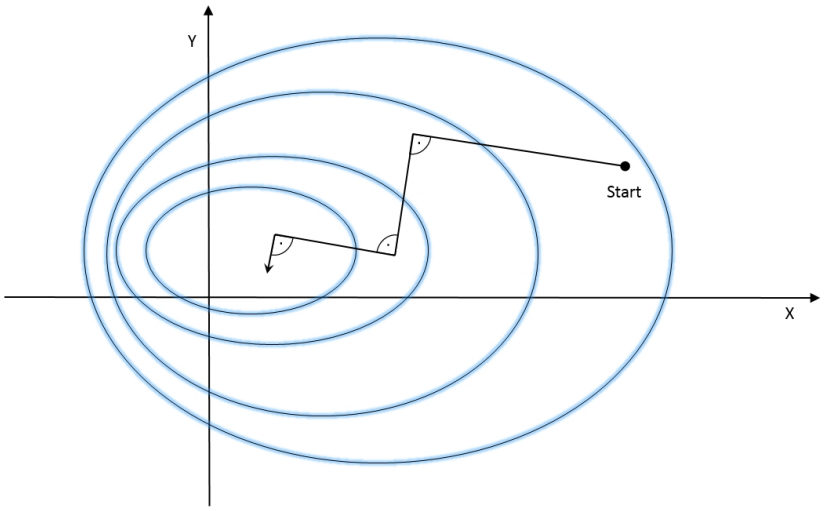


Abbildung 2.3: *Die* Steepest-Descent-*Methode nähert sich dem Minimum in einem Zick-Zack-Muster, wobei jede neue Richtung orthogonal zur vorherigen Richtung verläuft*[103].

Conjugate-Gradient (CG)

Das SD-Verfahren konvergiert nur langsam in der Nähe des Minimums, weil neue Segmente des Pfades den Fortschritt früherer Iteration zunichte machen können. Wie in Abbildung 2.4 zusehen ist, weicht der Pfad der SD-Methode (schwarz) von der idealen Richtung zum Minimum ab, da jede neue Richtung orthogonal zur vorangehenden ist. Die *Conjugate-Gradient*-(CG)-Methode ist ein Minimierungsverfahren zur Lösung von großen linearen Gleichungssystemen. Das CG-Verfahren korrigiert kontinuierlich die Richtung des Gradientenpfades[103].

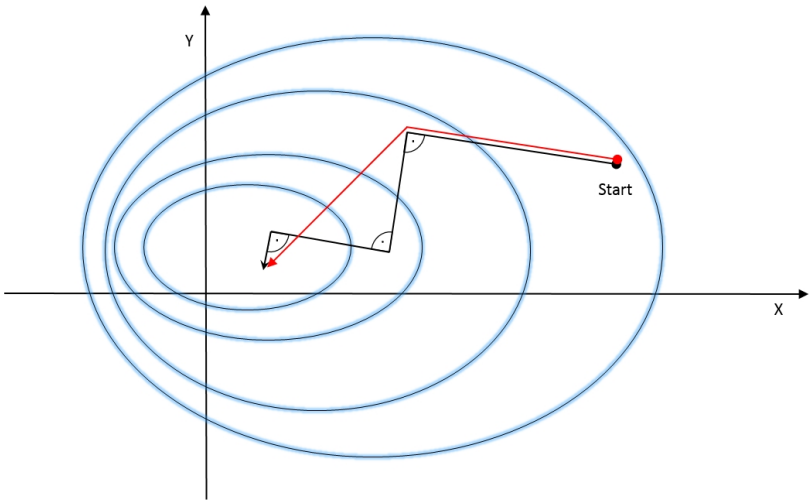


Abbildung 2.4: *Vergleich der* Steepest-Descent-*Methode (schwarz) mit der* Conjugate-Gradient-*Methode (rot). Während das SD-Verfahren sich im Zick-Zack-Muster dem Minimum nähert, optimiert die CG-Methode die Richtung bestmöglich.*

Wie auch beim SD-Verfahren ist das folgende Minimierungsproblem ohne Nebenbedingungen gegeben:

$$\min \ f(x), \quad mit \ x \in R^n \tag{2.3}$$

wobei R^n ein n-dimensionaler euklidischer Raum ist und $f : R^n \to R$ eine kontinuierlich differenzierbare Funktion darstellt. Die CG-Methode hat dann die Form:

$$x_{k+1} = x_k + \alpha_k d_k, \quad mit \ k = 0, 1, 2, \ldots \tag{2.4}$$

wobei x_0 der Startpunkt, a_k die Schrittweite und d_k als

$$d_k = \begin{cases} -g_k & , k = 0 \\ -g_k + \beta_k d_{k-1} & , k > 0 \end{cases} \tag{2.5}$$

gewählt werden kann. Dabei ist $g_k = \nabla f(x_k)$ der Gradient von $f(x_k)$ und β_k bestimmt die gewählte CG-Methode. In dieser Arbeit wird das Verfahren von

Polak-Ribiére-Polyak[104,105] verwendet. Es ist definiert als[106]:

$$\beta^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{||g_{k-1}||^2} \tag{2.6}$$

Alle klassischen gradient-basierten Minimierungsverfahren weisen dasselbe Problem auf: Sie garantieren nicht, überhaupt ein Minimum zu finden, sondern können sogar an Sattelpunkten und lokalen Maxima stehen bleiben. Die Gradientverfahren stoppen immer dann, wenn der Gradient verschwindet[107].

2.1.2 Molekulardynamik (MD)

Die Molekulardynamik (engl. *Molecular Dynamics*, kurz MD)[2–4] ist eine Methode für die Berechnung von thermodynamischen Eigenschaften eines klassischen Vielteilchensystems, dessen gewöhnliche Differentialgleichungen 2. Ordnung für mehr als zwei Teilchen nicht mehr durch algebraische Rechenoperationen lösbar sind (Satz von Bruns und Poincaré[1]). Hierbei werden die Anfangskoordinaten r_i und Geschwindigkeiten v_i der Atome anhand einer Serie von endlichen Zeitschritten fortgepflanzt, sodass eine Abfolge von zeitlich korrelierten Punkten im Phasenraum - eine Trajektorie - entsteht[101]. Die Dynamik der Atome gehorcht dabei der klassischen Newtonschen Mechanik bzw. dem zweiten Newtonschen Axiom:

$$F_i = \frac{dv_i}{dt} m_i, \; v_i = \frac{dr_i}{dt}, \tag{2.7}$$

wobei r_i die Raumkoordinate, v_i die Geschwindigkeit, m_i die Masse, F_i die Kraft auf Partikel (Atom) i und t die Zeit sind. Die auf die Atome wirkenden Kräfte berechnen sich anhand der parametrisierten Energiefunktionen des angewendeten Kraftfeldes (hier das COMPASS-Kraftfeld, siehe Kapitel 2.1.3). Die Integration der Bewegungsgleichungen erfolgt numerisch. Das in dieser Arbeit angewandte Integrationsschema basiert auf dem *Velocity-Verlet*-Algorithmus.

Eine MD-Simulation mit einer bestimmten Anzahl an Partikeln N, einem festen Volumen V und einer konservierten Gesamtenergie E läuft im mikrokanonischen Ensemble (NVE) ab. Diese Bedingungen entsprechen häufig nicht

denen, wie sie normalerweise in Experimenten vorliegen. Experimente finden häufig bei konstanter Temperatur oder konstantem Druck statt. Um eine konstante Temperatur T_{md} zu erzwingen, können die Geschwindigkeiten der Partikel v_t in jedem Zeitschritt t skaliert werden, damit die kinetische Energie des Systems als beständig angesehen werden kann. Dieser Ansatz wird als *Velocity-Scaling* bezeichnet[108] und ist definiert als:

$$v_{t+1} = \sqrt{\frac{T_{md}}{T(t)}} \cdot v_t \tag{2.8}$$

wobei v_{t+1} die resultierende Geschwindigkeit ist und $T(t)$ die aktuelle Temperatur des Systems zum Zeitpunkt t. Die Technik des *Velocity-Scaling* wird in dieser Arbeit verwendet, um im kanonischen Ensemble (NVT) die MD-Simulation für die Berechnung der paarweisen Wechselwirkungsenergien E_{ij} (siehe Kapitel 3.1.3) durchführen zu können[108–112].

Integrationsschema Velocity-Verlet

Für die Lösung des Vielteilchenproblems sind numerische Approximationsverfahren notwendig. In den 1960er entwickelte Loup Verlet[113] ein Integrationsschema zur numerischen Lösung der Newtonschen Bewegungsgleichungen für die Anwendung in MD-Simulationen. Der Algorithmus ist nicht nur einer der am weitesten verbreitetsten Techniken[114], sondern auch eines der einfachsten und dabei genauesten Verfahren[33]. Aufgrund der Zeitreversibilität des Verlet-Algorithmus ist eine hohe Energiekonservierung während der Simulationen gegeben, sodass der Verlet-Algorithmus anderen Verfahren wie den Runga-Kutta-Methoden vorzuziehen ist[101].

Schreibt man das zweite Newtonsche Axiom $F = ma$ in seiner differentiellen Form auf:

$$-\frac{dV}{dr} = m\frac{d^2r}{dt^2} \tag{2.9}$$

wobei V die potentielle Energie an Position r ist und der Vektor r alle Koordinaten

aller Partikel in dem betrachteten System beinhaltet. Es besteht die Möglichkeit, für diese Partikel die Positionen einen kleinen Zeitschritt Δt später mittels einer Taylorreihenentwicklung zu bestimmen:

$$\begin{aligned} r_{i+1} &= r_i + \frac{\delta r}{\delta t}(\Delta t) + \frac{1}{2}\frac{\delta^2 r}{\delta t^2}(\Delta t)^2 + \frac{1}{6}\frac{\delta^3 r}{\delta t^3}(\Delta t)^3 + \dots \\ r_{i+1} &= r_i + v_i(\Delta t) + \frac{1}{2}a_i(\Delta t)^2 + \frac{1}{6}b_i(\Delta t)^3 + \dots \end{aligned} \tag{2.10}$$

wobei v_i die Geschwindigkeit, a_i die Beschleunigung und b_i die Hyperbeschleunigung darstellen.

Die Positionen der Partikel zu einem Zeitschritt Δt früher kann aus Gleichung 2.10 durch Substitution von Δt mit $-\Delta t$ hergeleitet werden:

$$r_{i-1} = r_i - v_i(\Delta t) + \frac{1}{2}a_i(\Delta t)^2 - \frac{1}{6}b_i(\Delta t)^3 + \dots \tag{2.11}$$

Die Addition der Gleichungen 2.11 und 2.10 liefert den *Verlet*-Algorithmus in seiner Grundform. Dieser berechnet aus der momentanen Position r_i, der vorherigen Position r_{i-1} und der momentanen Beschleunigung des Partikels a_i die Position des fokussierten Partikels einen Zeitschritt Δt später. Die Lösung der Bewegungsgleichungen mit dem *Verlet*-Algorithmus ist somit ein iterativer Prozess mit der Schrittweite Δt[101]:

$$r_{i+1} = (2r_i - r_{i-1}) + a_i(\Delta t)^2 \tag{2.12}$$

Die Formulierung des ursprünglichen *Verlet*-Algorithmus (Gleichung 2.12) bringt ein Problem mit sich. Aufgrund der Addition der Terme 2.11 und 2.10 fällt die Geschwindigkeit v aus der Berechnung heraus. Die Geschwindigkeit v_i ist nicht direkt berechenbar, da r_{i+1} bekannt sein muss:

$$v_i = \frac{r_{i+1} - r_{i-1}}{2\Delta t} \tag{2.13}$$

Daraus folgt, dass die berechnete Geschwindigkeit immer einen Zeitschritt gegenüber der Berechnung der Raumkoordinaten zurückliegt[114]. Die kinetische Energie des betrachteten Systems ist jedoch eine wichtige Kenngröße. Die Lösung für

dieses Problem ist der *Velocity-Verlet*-Algorithmus[115], welcher die Berechnung der Raumkoordinaten r und der Geschwindigkeit v phasengleich ermöglicht[101]:

$$r_{i+1} = r_i + v_i \Delta t + \frac{1}{2} a_i \Delta t^2 \tag{2.14}$$

$$v_{i+1} = v_i + \frac{1}{2}(a_i + a_{i+1})\Delta t \tag{2.15}$$

2.1.3 Das COMPASS-Kraftfeld

Die Wahl des richtigen Kraftfeldes für die Bestimmung des DPD-Repulsionsparameters a_{ij} ist essentiell. Da die dissipative Partikeldynamik für die Berechnung von Eigenschaften komplexer Flüssigkeiten gestaltet wurde[34,35], sollte das zugrunde liegende Kraftfeld für die Simulation von kondensierten Phasen optimiert sein. Insbesondere die nicht-bindenden Wechselwirkungen des Kraftfeldes sind interessant. Die Berechnung des DPD-Repulsionsparameters basiert größtenteils auf den nicht-bindenden Anteilen.

In den letzten Jahrzehnten galt der Entwicklung von Kraftfeldern zur Bestimmung von Eigenschaften kondensierter Phasen große Aufmerksamkeit. Dies zeigt deutlich die Arbeit an Kraftfeldern wie AMBER[116] (Assisted Model Building with Energy Refinement) und CHARMM[117,118] (Chemistry at HARvard Macromolecular Mechanics), welche speziell für Aufgaben in der molekularen Biologie konstruiert werden[100].

Im Rahmen dieser Arbeit wird das COMPASS (*Condensed-Phase Optimized Molecular Potentials for Atomistic Simulations Studies*) Kraftfeld verwendet. Das COMPASS-Kraftfeld wurde speziell konstruiert, um kondensierte Phasen und die korrekte Beschreibung derer nicht-bindender Wechselwirkungen wiederzugeben. Es ist ein Klasse 2-Kraftfeld mit zusätzlichen Kopplungstermen. Kopplungsterme beschreiben Abhängigkeiten verschiedener Valenzterme untereinander, d.h. bei Änderungen einer Bindungslänge resultiert dies beispielsweise in einer gleichzeitigen Anpassung eines Bindungswinkels oder eines Torsionswinkels. Die Parametrisierung erfolgte unter Verwendung von *ab-initio*-Berechnungen und empirischen Daten. Die Partialladungen wurden aus *ab-initio*-Berechnung gewonnen, die

Valenzparameter teilweise vom PCFF-Kraftfeld (*Polymer Consistent Force Field*) übernommen bzw. wenn nicht vorhanden, aus *ab-initio*-Verfahren berechnet. Das intermediäre Kraftfeld wird anschließend anhand von empirischen Daten validiert und verfeinert. Gleichung 2.16 zeigt die funktionale Form des COMPASS-Kraftfeldes:

$$
\begin{aligned}
E_{total} = & \underbrace{\sum_{b}[k_2(b-b_0)^2 + k_3(b-b_0)^3 + k_4(b-b_0)^4]}_{1} + \\
& \underbrace{\sum_{\theta}[k_2(\theta-\theta_0)^2 + k_3(\theta-\theta_0)^3 + k_4(\theta-\theta_0)^4]}_{2} + \\
& \underbrace{\sum_{\phi}[k_1(1-\cos\phi) + k_2(1-\cos 2\phi) + k_3(1-\cos 3\phi)]}_{3} + \\
& \underbrace{\sum_{\chi} k_2\chi^2}_{4} + \underbrace{\sum_{b,b'} k(b-b_0)(b'-b'_0)}_{5} + \underbrace{\sum_{b,\theta} k(b-b_0)(\theta-\theta_0)}_{6} + \\
& \underbrace{\sum_{b,\phi}(b-b_0)[k_1\cos\phi + k_2\cos 2\phi + k_3\cos 3\phi]}_{7} + \\
& \underbrace{\sum_{\theta,\phi}[k(\theta-\theta_0)(\theta'-\theta'_0)(k_1\cos\phi + k_2\cos 2\phi + k_3\cos 3\phi)]}_{8} + \\
& \underbrace{\sum_{\theta,\phi}[k(\theta-\theta_0)(\theta'-\theta'_0)\cos\phi]}_{9} + \\
& \underbrace{\sum_{i,j}\frac{q_i q_j}{r_{ij}}}_{10} + \underbrace{\sum_{i,j} E_{ij}\left[2\left(\frac{r^0_{ij}}{r_{ij}}\right)^9 - 3\left(\frac{r^0_{ij}}{r_{ij}}\right)^6\right]}_{11}
\end{aligned}
\tag{2.16}
$$

Die einzelnen Terme können in zwei Gruppen eingeteilt werden:

- Valenzterme (1-9):
  Bindungen (b), Bindungswinkel (θ), Torsionswinkel (ϕ), *Out-of-Plane*-Winkel (χ) und die zugehörigen Kopplungsterme.

- Nicht-bindende Terme (10,11):

Coulombfunktion für elektrostatische Wechselwirkungen und eine 9-6-Lennard-Jones-(LJ)-Potentialfunktion für die Beschreibung der Van-der-Waals-Wechselwirkung (VdW).

Im Gegensatz zur Beschreibung der VdW-Wechselwirkungen mit dem 12-6-LJ-Potential, welches zu *hart* im repulsiven Bereich ist, verwendet das COMPASS-Kraftfeld ein 9-6-LJ-Potential. Dieses ist *weicher* im repulsiven Anteil, aber mit der Konsequenz gelegentlich zu attraktiv im langreichweitigen Anteil zu sein[119]. Abbildung 2.5 zeigt beide LJ-Potentiale im Vergleich.

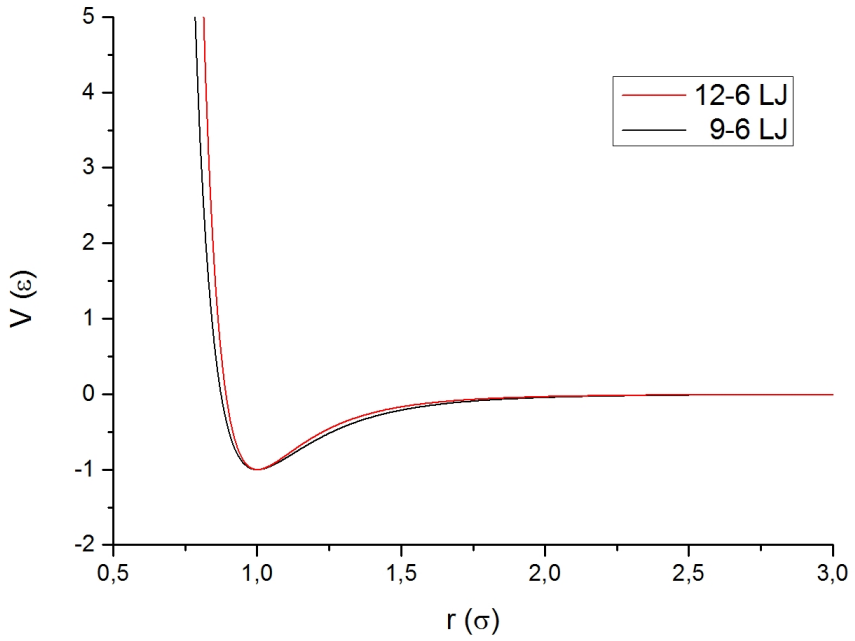


Abbildung 2.5: *Vergleich des 12-6-LJ-Potentials mit dem 9-6-LJ-Potential.*

Die elektrostatische Wechselwirkung wird über Partialladungen der Atome dargestellt. Damit die Ladungsparameter auf andere Systeme transferierbar bleiben, werden sie als sogenannte *Bond-charge increments* (BCIs) δ_{ij} festgehalten. Diese repräsentieren die Ladungsseperation zwischen zwei valenzgebundenen Atomen i und j. Die Partialladung für Atom i ist somit:

$$q_i = \sum_j \delta_{ij} \tag{2.17}$$

wobei j alle Atome beschreibt, welche valenzgebunden mit Atom i sind.

Die Validierung des Kraftfeldes wurde anhand von 172 Molekülen, 102 Flüssigkeiten und 69 Kristallen durchgeführt, welche 28 der häufigsten molekularen Klassen repräsentieren. Die berechneten Strukturparameter (Bindungslänge und Winkel) weichen dabei ungefähr 1.0 % von den experimentellen Daten ab. Die vorhergesagten Parameter für Flüssigkeiten zeigen eine mittlere quadratische Abweichung für die Dichte von 1.9 % und für die kohäsiven Energien von 4.1 %[100].

2.1.4 Dissipative Partikeldynamik (DPD)

Trotz Vereinfachung der Modelle auf ein atomistisches Niveau sind die Kraftfeldmethoden immer noch zu feinkörnig (engl. *fine-grained*), als dass man Systeme mit mehreren Millionen Atomen, Raumskalen von einigen Mikrometern sowie zeitlichen Auflösungen von mehreren Mikrosekunden in angemessener Zeit berechnen könnte. Die dissipative Partikeldynamik ist eine mesoskopische Simulationstechnik, welche die Betrachtung von kondensierter Materie in sowohl atomistischen, als auch mesoskopischen Größen ermöglicht. Die DPD-Methode vergröbert (engl. *coarse-graining*) den MM-Ansatz, indem sogenannte *Beads* eingeführt werden und die bisherigen Atomtypen ersetzen. Diese *Beads* können aus einem Atom, einem molekularem Fragment, einem ganzen Molekül und sogar aus einer beliebigen Raumregion einer Flüssigkeit bestehen[33,120].

Die dissipative Partikeldynamik wurde ursprünglich von Koelman und Hoogerbrugge[34,35] in den frühen 1990er Jahren entwickelt und basiert auf der Molekulardynamik. Die DPD-*Beads* sind paarweise miteinander wechselwirkende chemische Entitäten, welche sich in diskreten zeitlichen Schritten durch den Raum bewegen. Die Bewegung wird anhand des zweiten Newtonschen Axioms beschrieben:

$$F_i = \frac{dv_i}{dt} m_i, \; v_i = \frac{dr_i}{dt}, \tag{2.18}$$

wobei r_i die Raumkoordinate, v_i die Geschwindigkeit, m_i die Masse und F_i die

gesamte Kraft ist, die auf ein *Bead* wirkt. t stellt die Zeit dar. Auch an dieser Stelle ist die Lösung der Newtonschen Bewegungsgleichungen (Vielteilchenproblem) essentiell. Erste Ansätze nutzten einen Algorithmus basierend auf dem Eulerverfahren für die Lösung der Differentialgleichungssysteme. Dies hatte zur Folge, dass die Temperatur des DPD-Systems immer höher lag als die thermodynamische Temperatur[121]. Für eine angemessenere Beschreibung der Temperatur schlugen Groot und Warren[120] vor, einen modifizierten *Velocity-Verlet*-Algorithmus zu verwenden.

Neben den konservativen Kräften, welche schon in MD-Techniken eingeführt wurden, interagieren zwei neue Kräfte mit den DPD-*Beads*: Die dissipative und die stochastische Kraft. Gleichung 2.19 zeigt die Gesamtkraft F_i, die auf ein *Bead* einwirkt:

$$F_i = \sum_{i \neq j} \left(F_{ij}^C + F_{ij}^D + F_{ij}^R \right), \tag{2.19}$$

wobei die Summe paarweise über alle *Beads* innerhalb eines Radius mit dem Grenzwert r_c berechnet wird. Die konservative Kraft spiegelt das Potential eines *Beadpaares* wieder, welches in die Richtung ihrer jeweiligen Zentren wirkt, d.h. die beiden interagierenden *Beads* verhalten sich repulsiv zueinander. Folgende Gleichung beschreibt die konservative Kraft:

$$F_{ij}^C = \begin{cases} a_{ij}(1 - r_{ij})\hat{r}_{ij} & (r_{ij} < r_c) \\ 0 & (r_{ij} \geq r_c) \end{cases}, \tag{2.20}$$

wobei a_{ij} der DPD-Repulsionsparameter zwischen *Bead* i und j ist, mit $r_{ij} = r_i - r_j$, $r_{ij} = |r_{ij}|$, $\hat{r}_{ij} = r_{ij}/|r_{ij}|$. Die zweite Kraft ist die dissipative Kraft, welche als eine Reibungskraft interpretiert werden kann. Sie ist abhängig von der Position r_i und der Geschwindigkeit v_i der Partikel:

$$F_i^D j = -\gamma_{DPD} \omega^D(r_{ij})(\hat{r}_{ij} \cdot v_{ij})\hat{r}_{ij} \tag{2.21}$$

wobei der Reibungskoeffizient γ_{DPD} die Stärke der Reibung und $\omega^D(r_{ij})$ die Größe des Einfluss der Distanz der *Beads* zueinander beschreibt. Die dritte Kraft ist die

stochastische Kraft:

$$F_i^R j = \sigma \omega^R(r_{ij}) \theta_{ij} \hat{r}_{ij} \tag{2.22}$$

wobei σ die Stärke der stochastischen Kraft kontrolliert, $\omega^R(r_{ij})$ wiederum den Einfluss der Distanz der *Beads* zueinander skaliert, und θ_{ij} eine standardverteilte Zufallsvariable mit Einheitsvarianz ist.[33,38,39,120]

Die Funktionen $\omega^D(r_{ij})$, $\omega^R(r_{ij})$ und die Koeffizienten γ, σ können nicht unabhängig voneinander gewählt werden. Español und Warren[39] haben gezeigt, dass DPD-Simulationen nur im NVT-Ensemble ablaufen, wenn die Gewichtungsfunktionen wie folgt definiert werden:

$$\gamma_{DPD} = \frac{\sigma^2}{2k_B T} \tag{2.23}$$

$$\omega^D(r_{ij}) = \left[\omega^R(r_{ij})\right]^2 \tag{2.24}$$

wobei k_B die Boltzmann-Konstante und T die Temperatur ist. Diese beiden Funktionen beschreiben das sogenannte Fluktuierende-Zerstreuungs-(engl. *fluctuation-dissipation*)-Theorem der dissipativen Partikeldynamik.

2.1.5 Molekulare Fragmentdynamik (MFD)

Die molekulare Fragmentdynamik (MFD) ist eine chemisch intuitive Variante der dissipativen Partikeldynamik. Die *Beads* sind so gewählt, dass sie explizit konkrete Moleküle oder Molekülfragmente des zu untersuchenden chemischen Ensembles darstellen. Die Gesamtheit aller Molekülfragmente einer MFD-Simulation ist das Fragmentset. Jedes Molekül im System wird aus diesem Fragmentset aufgebaut. Die einzelnen Fragmente werden mittels harmonischer Federn zur Darstellung kovalenter Bindungen verknüpft. Die Fragmenttopologie eines Moleküls wird als dessen Fragmentierungsschema bezeichnet[48,56,57].

Die im Rahmen von dieser Arbeit verwendeten DPD-Repulsionsparameter des hier zu entwickelten Fragmentsets werden anhand von MD-Simulationen unter

Verwendung des COMPASS-Kraftfeldes berechnet.

Simulationssoftware MFD-FormulaOne

MFD-FormulaOne[54] ist eine umfassende *Rich-Client*-Anwendung für mesoskopische Simulationen mit der molekularen Fragmentdynamik. Es wird in Kooperation von CAM-D Technologies und GNWI - Gesellschaft für naturwissenschaftliche Informatik entwickelt. Die Anwendung unterstützt den kompletten Simulationsprozess beginnend beim Design über die Ausführung bis hin zur Analyse von MFD-Simulationen (Jobs). Für alle Aufgaben, von der Erstellung bis hin zur Analyse, bietet MFD-FormulaOne eine vollständig integrierte Benutzeroberfläche.

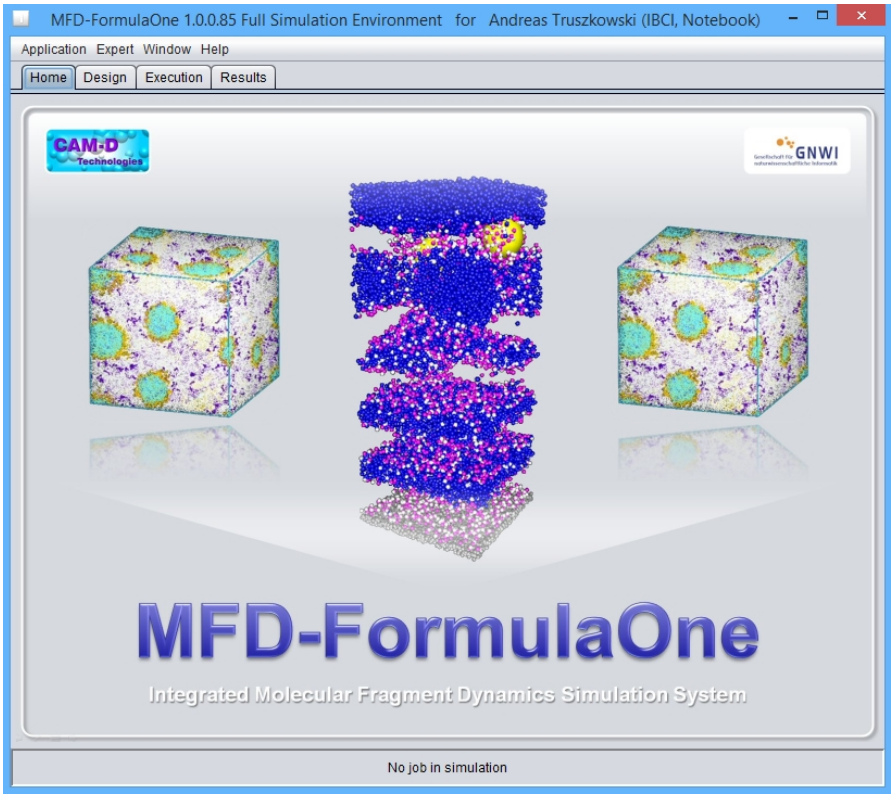


Abbildung 2.6: *Startbildschirm der Anwendung MFD-FormualOne*[54].

Das Programm gliedert sich in vier Bereiche:

- **Jobdesign:** Die Jobdesignoberfläche listet alle verfügbaren Simulationsjobs auf, wobei auch Filter angewendet werden können. Die Oberfläche erlaubt

das Verändern, Wiederverwenden, Erstellen oder auch das Archivieren von verfügbaren Jobs. Die Beschreibung eines Jobs gliedert sich in vier Sektionen: (1) Allgemeine Beschreibung, (2) Beschreibung des chemischen Systems, (3) Beschreibung der Wechselwirkungen und (4) die Simulationsparameter. Jede Sektion besteht aus mehreren Simulationseigenschaften, welche alle mit einem Hilfetext versehen sind.

- **Jobausführung:** MFD-FormulaOne erlaubt die parallelisierte Ausführung von unterschiedlichen Jobs in Abhängigkeit der Einstellungen und der Anzahl der verfügbaren Prozessorkerne der verwendeten Hardware. Der Fortschritt einer Simulation lässt sich während der Ausführung betrachten. Zusätzlich gibt es die Möglichkeit, bereits abgeschlossene Berechnungen mit einer bestimmten Anzahl an Iterationsschritten weiter rechnen zu lassen.

- **Ergebnisanalyse:** Für die Analyse von Simulationsergebnissen (z.B. Längenmessungen) bietet MFD-FormulaOne Simulationsboxansichten und unterstützt die Erzeugung von Filmen einer Simulationsdynamik. Eigenschaften wie die radiale Verteilungsfunktionen und Distanzen von Fragmentpaaren, Oberflächenspannung, molekularer Gyrationsradius, usw. werden zusammengefasst und in Diagrammen dargestellt.

- **Jobmanagment:** Jobs können verschoben oder als einzelne Dateien komprimiert werden. Verzeichnisse mit beinhalteten Simulationen können einfach synchronisiert werden, sodass das Jobdesign auf einem Notebook vorgenommen werden kann und anschließend die Daten zur Ausführung auf einen Server transferiert werden können.

Berechnung des DPD-Repulsionsparameters

Die *Beads* in der molekularen Fragmentdynamik repräsentieren explizit konkrete Moleküle oder Molekülfragmente. Nach Brini et al.[49] ist es möglich, die Paarpotentiale der Partikel aus *fine-grained* Simulationen zu berechnen (*bottom-up*-Verfahren), d.h. aus den direkten Wechselwirkungen der Moleküle in einer Computersimulation wie MC- oder MD-Simulationen. Groot und Warren[120] haben

gezeigt, dass der DPD-Repulsionsparamter a_{ij} in einem direkten linearen Zusammenhang mit dem Flory-Huggins-Parameter χ_{ij} steht:

$$a_{ij}(T) = 25 + 3.497 \cdot \chi_{ij}(T) \tag{2.25}$$

Paul Flory [122] und Maurice Huggins [123] entwickelten unabhängig voneinander ein einfaches Gittermodell zur Beschreibung von nicht-idealen Polymerlösungen. Das Flory-Huggins-Modell zeigt, dass der Flory-Huggins-Parameter χ_{ij} eine dimensionslose Größe ist, welche die Interaktionsenergie eines Partikels mit einem Lösungsmittel beschreibt:

$$\chi_{ij} = \frac{Z_{ij} \cdot \Delta E_{ij}}{RT} \tag{2.26}$$

wobei Z_{ij} die Koordinationszahl des untersuchten Molekülpaares, ΔE_{ij} die Energiedifferenz der paarweisen Wechselwirkung, R die universelle Gaskonstante und T die Temperatur sind. Die Bestimmung der Energiedifferenz der paarweisen Wechselwirkung ΔE_{ij} kann empirisch erfolgen (Verdampfungsenthalpie) oder mittels computergestützte Verfahren ermittelt werden. Diese können Methoden des maschinellen Lernens in Kombination mit Quantitativer Struktur-Wirkungs-Beziehung (engl. *Quantitative Structure-Activity Relationship*, kurz QSAR) sein sowie anhand von atomistischen Verfahren wie MD- oder MC-Simulationen kalkuliert werden. Die Berechnung der Energiedifferenz der paarweisen Wechselwirkungen ΔE_{ij} erfolgt anhand der folgenden Formel [124]:

$$\Delta E_{ij} = \frac{1}{2}(E_{ij} + E_{ji}) - \frac{1}{2}(E_{ii} + E_{jj}) \tag{2.27}$$

Im Rahmen dieser Arbeit wird die Energiedifferenz der paarweisen Wechselwirkung ΔE_{ij} anhand von atomistischen MD-Simulationen mit der Parametrisierung des COMPASS-Kraftfeldes ermittelt. Hierbei wird die freie Mischungsenergie der Moleküle A und B eines Molekülpaares bestimmt. Der χ-Parameter repräsentiert die Repulsionsenergie der beiden Komponenten zueinander, also die Stärke der gemittelten Abstoßung von Molekül A zu B und umgekehrt [48,50–53]. Abbildung 2.7 skizziert die paarweisen Weschelwirkungsenergien eines Polymers A und dem

Lösungsmittel B in einem Gittermodell.

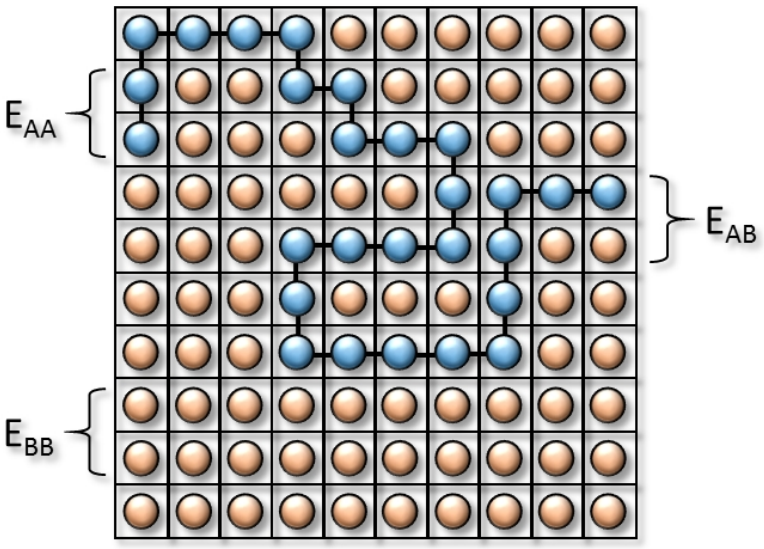


Abbildung 2.7: *Darstellung der paarweisen Wechselwirkungsenergien E_{ij} einer Polymerlösung in einem Gittermodell (Polymer A: Blau; Lösungsmittel B: Orange).*

Die Methode für die Berechnung des DPD-Repulsionsparameters a_{ij} ist in Kapitel 3.1.3 beschrieben. Aus Gleichung 2.26 und 2.27 folgt, dass die paarweisen Wechselwirkungsenergien E_{AA}, E_{BB}, E_{BA} und E_{AB} sowie die Koordinationszahl Z_{AB} für jedes Molekülpaar eines Fragmentsets zur Bestimmung des Repulsionsparameters a_{ij} ermittelt werden müssen.

2.1.6 Weitere mesoskopische Simulationstechniken

Neben der hier verwendeten molekularen Fragmentdynamik auf Basis der dissipativen Partikeldynamik gibt es noch weitere mesoskopische Simulationstechniken, welche die Dynamik von Molekülen in kondensierten Phasen betrachten. Alle hier vorgestellten Simulationstechniken weisen eine Gemeinsamkeit auf: Eine mit geringem Rechenaufwand verbundene Betrachtung von mesoskopischen Flüssigkeiten. Zwei weitere mesoskopische Methoden die auch auf Partikeln basieren, sind die *Lattice-Boltzmann*-(LB)-[125–127]Methode und die *Brownian-Dynamics* -

(BD)[128,129]-Methode[33].

Lattice-Boltzmann-(LB)-Methode

Wie dem Namen zu entnehmen ist, basiert die *Lattice-Boltzmann*-(LB)-Methode[125–127] auf einem Gitterschemata und hat ihren Ursprung in Flüssigkeitsmodellen von zellularen Gitter-Gas-Automaten[130]. Anstatt die Navier-Stokes-Gleichungen zur Beschreibung der Strömung von Flüssigkeiten zu lösen, nutzt diese Methode die numerische Lösung der Boltzmann-Gleichung für die Simulation von Flüssigkeiten:

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial r} + \frac{F}{m} \cdot \frac{\partial f}{\partial v} = \left(\frac{\partial f}{\partial t}\right)_{St} \tag{2.28}$$

wobei $\left(\frac{\partial f}{\partial t}\right)_{St}$ der Stoßterm, r der Ort und v die Geschwindigkeit eines Teilchens zur Zeit t der Masse m unter dem Einfluss der äußeren Kraft F sind. $f(r, v, t)$ stellt die Verteilungsfunktion im Phasenraum dar.

In einem LB Modell wird die Flüssigkeit über ein reguläres Gitter dargestellt, wobei jeder Gitterpunkt z nächste Nachbarn hat. Neben der räumlichen Diskretisierung zur Lösung der Boltzmann-Gleichung ist die Zeit ebenfalls diskretisiert. Der Algorithmus läuft in zwei Schritten ab: Strömung und Kollision. Im ersten Schritt der Simulation bewegt sich ein Teilchen entlang der Verbindung zu einem benachbarten Gitterpunkt. Hierbei kann sich nur ein Teilchen zum selben Zeitpunkt von seinem aktuellen Gitterpunkt über die Verbindung zum benachbarten Gitterpunkt bewegen. Alle Teilchen bewegen sich hierbei gleichzeitig. Der zweite Simulationsschritt ist der Kollisionsschritt. Dieser definiert Regeln zur Einhaltung der maximalen Anzahl an Teilchen sowie des Gesamtimpulses eines Gitterpunktes. Abgesehen von diesen Einschränkungen kann jedes Teilchen seine Geschwindigkeit ändern. Der entscheidende Nachteil dieser Methode ist die Notwendigkeit eines einfach strukturiertem Gitters. Die Verwendung anderer Gitterstrukturen ist nur bedingt möglich[33].

Brownian Dynamics-(BD)-Methode

Die *Brownian Dynamics*-(BD)-Methode[128,129] dient der Beschreibung von Teilchen, die der Brownschen Molekularbewegung unterliegen und bildet die Basis der dissipativen Partikeldynamik. Der wesentliche Unterschied zwischen der dissipativen Partikeldynamik und *Brownian Dynamics* ist, dass die BD-Methode nicht das dritte Newtonsche Axiom erfüllt und somit nicht dem Impulserhaltungssatz unterliegt. Folglich reproduziert die BD-Methode niemals hydrodynamisches Verhalten[33]. Diesen Unterschied macht eine Untersuchung von Groot et al.[131] zum Phasenverhalten von Diblockcopolymeren deutlich. DPD-Simulationen entwickeln sich schnell zu den in Experimenten beschriebenen hexagonalen Strukturen, wobei die BD-Kalkulationen in einem metastabilen Zustand verbleiben.

Der $i$te Freiheitsgrad eines Teilchens wird mittels der Langevin-Gleichung wie folgt beschrieben:

$$\dot{v}_i(t) = a_i(t) - \gamma_i v_i(t) + \eta_i(t), \text{ mit } \dot{x}_i(t) = v_i(t) \tag{2.29}$$

wobei $a_i(t)$ die Beschleunigung des $i$ten Freiheitsgrades ist,

$$a_i(t) = \frac{F_i\left(\{r(t)\}\right)}{m_i} \tag{2.30}$$

wobei x_i der Ort, v_i die Geschwindigkeit und $\eta_i(t)$ eine Zufallsvariable mit einem Mittelwert von Null, keiner Korrelation mit x oder v in der Vergangenheit sowie in der Zukunft und mit der Autokorrelationsfunktion

$$\langle \eta_i(t)\eta_j(t+\Delta t)\rangle = 2\frac{k_B T_{ref}}{m_i}\gamma_i \delta(\Delta t)\delta_{ij} \tag{2.31}$$

γ_i ist hier die Reibungskonstante des $i$ten Freiheitsgrades[132,133]. Gleichung 2.29 ist das BD-Äquivalent zu Gleichung 2.19 der DPD-Methode.

2.2 Phospolipiddoppelschichten

In biologischen Systemen übernehmen Phospholipiddoppelschichten, auch als Biomembranen bezeichnet, die Rolle, unterschiedliche Kompartimente innerhalb oder zwischen einzelnen Zellen zu unterteilen und getrennte Reaktionsräume zu schaffen. Die einzelnen Bausteine dieser Membranen sind verschiedene Sorten der Phospholipide. Der Aufbau der Phospholipide (siehe Abbildung 2.8) besteht aus einer hydrophilen Kopfgruppe, einer Phosphatgruppe, als auch einem Glycerolrückgrat mit zwei abschließenden Fettsäuren. Die Fettsäuren können sowohl gesättigt als auch (cis-) ungesättigt vorliegen und unterschiedliche Kettenlängen aufweisen[75–78].

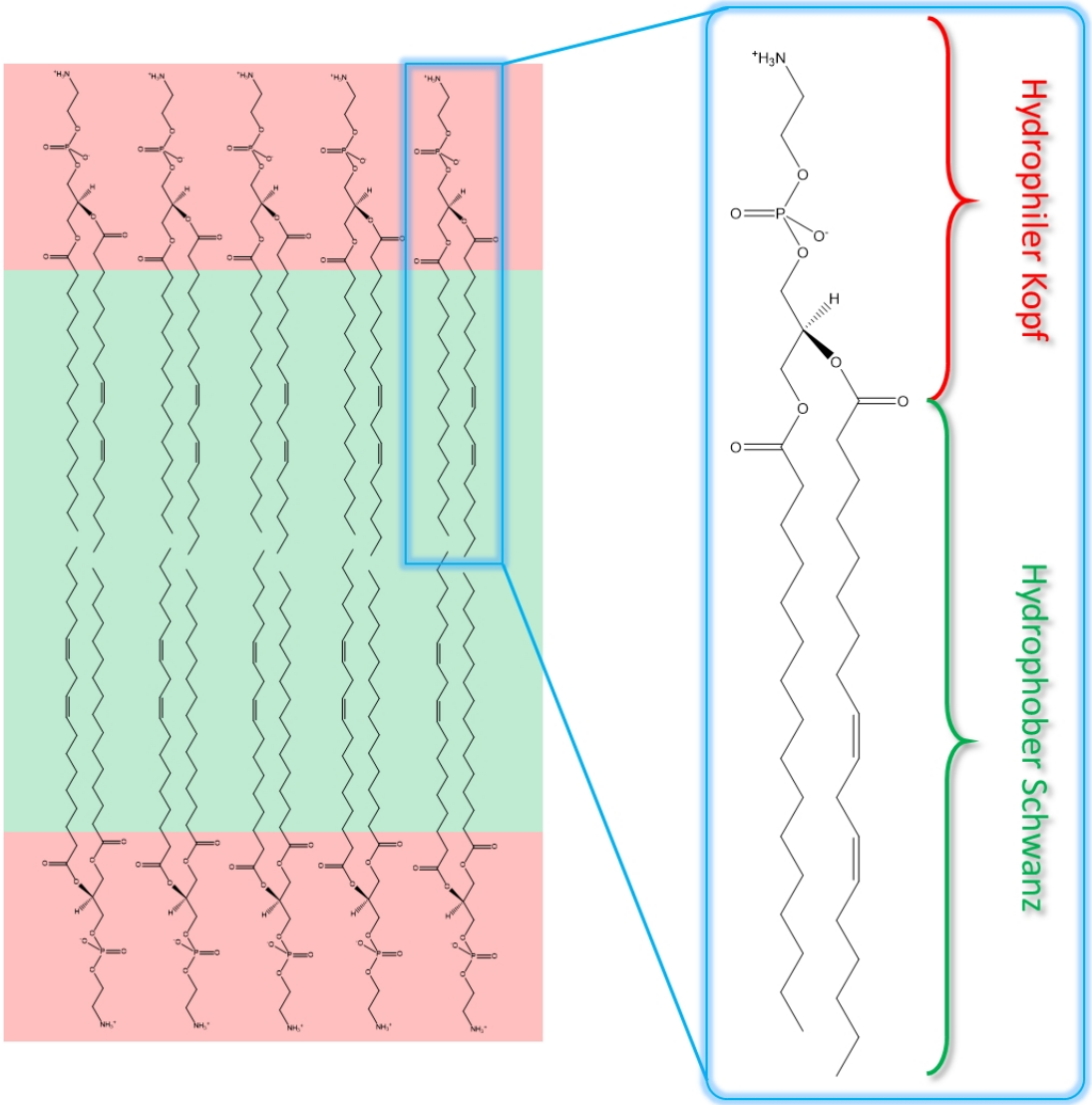


Abbildung 2.8: *Links: Schematischer Aufbau einer Phospholipiddoppelschicht. Rechts: Vergrößerte Darstellung des Phospholipids Phosphatidylethanolamin (PE).*

Die Ausbildung von Biomembranen durch Phospholipide innerhalb lebender Zellen basiert auf dem physikalischen Effekt, dass hydrophile Kopfgruppen dazu tendieren, sich in Richtung des wässrigen Mediums zu orientieren. Die hydrophoben Schwänze werden aufgrund der Entropie durch Wasser zur Selbstaggregation bewegt[76]. Hierdurch bildet sich eine stabile Phospholipiddoppelschicht aus, wobei die Phospholipide entgegengesetzt gegenüberliegend ausgerichtet sind. Die Phospholipide stehen nur durch ihre hydrophoben Hälften in Kontakt (siehe Abbildungen 2.8 und 2.9). Die Phospholipiddoppelschicht hat eine durchschnittliche

hydrophobe Dicke von 30 Å [76,83] und eine gesamte Dicke von 40-50 Å [84,85]. Innerhalb der Zellen ist die Zusammensetzung von Membranen vielfältig. Die Zusammensetzung der Membran des endoplasmatischen Retikulums (ER) und der Mitochondrienmembran sind einzigartig. Die häufigsten Lipide sind die Phosphoglyceride: Phosphatidylcholin (PC), Phosphatidylethanolamin (PE), Phosphatidylinositol (PI) und Phosphatidylserin (PS) [75,76,78]. Mit einem Anteil von mehr als 50 % ist Phosphatidylcholin das am häufigsten auftretende Phospholipid in eukaryotischen Membranen [78].

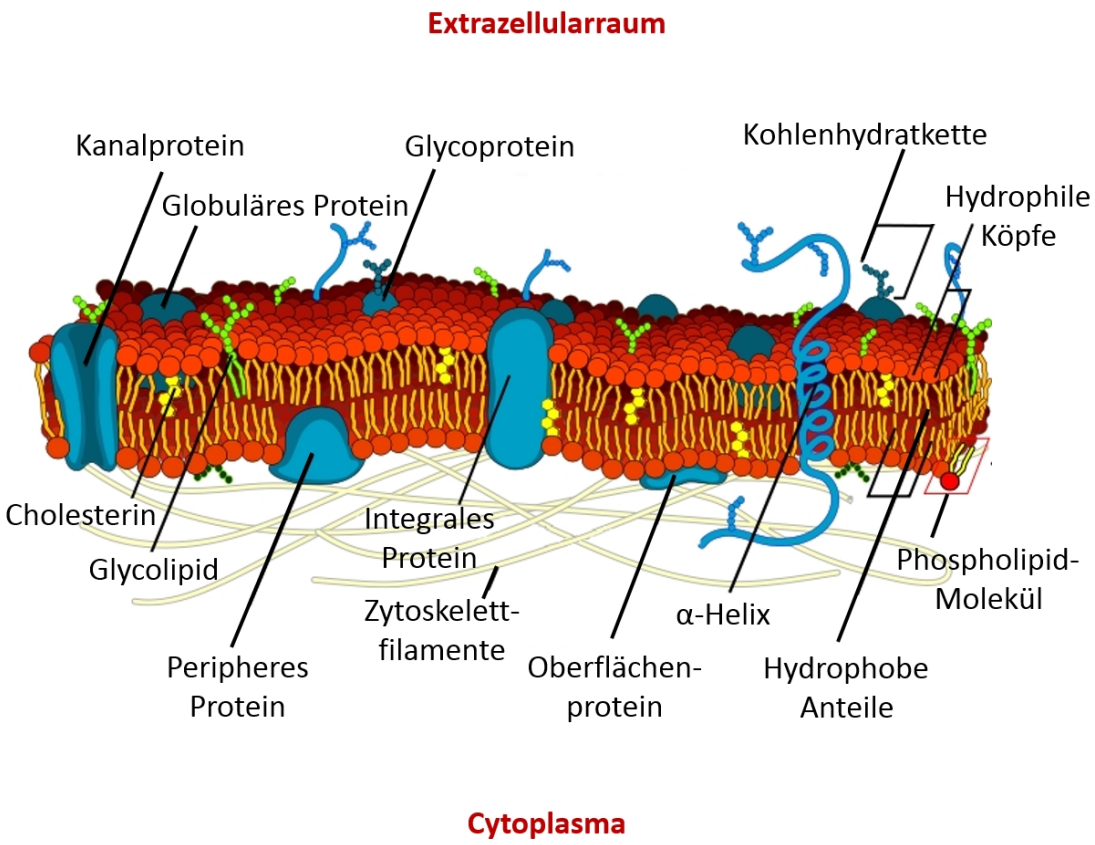


Abbildung 2.9: *Modell einer biologischen Membran.* [134].

Neben der Bereitstellung eines definierten Bereichs für spezifische biochemische Reaktionen innerhalb einer Zelle, besitzen Membranen vielfache Aufgaben. Nach dem Flüssig-Mosaik-Modell von 1972 übernehmen Phospholipiddoppelschichten Aufgaben eines überwiegend flüssigen Gerüstes bzw. einer Matrix für Membranproteine, welche lateral ungehindert durch die Membran diffun-

dieren können[135]. Hierbei unterstützen die vielseitigen Membrantypen eines biologischen Systems wichtige Prozesse einer Zelle wie Stofftransport, Sekretion, Energie- und Signaltransduktion, Zell-Zell-Kontakt und Proteinerkennung oder -transport[76,136]. Zusätzlich besitzen Phospholipiddoppelschichten weitere charakteristische Eigenschaften. Hierzu gehören die Bildung von Mizellen, Liposomen und Vesikeln, welche als zelluläre Transportmechanismen im Einsatz sind. Die Membranfusion von Phospholipiddoppelschichten ist eine zusätzliche mit Vesikelformierung zusammenhängende Fähigkeit, die eine wichtige Rolle für den zellulären Transport spielt[137,138]. Der Austausch von Lipiden zwischen den einzelnen Membranschichten wird als *Flip-Flops* bezeichnet[139,140].

2.3 Proteine

Proteine sind die vielseitigsten Makromoleküle in lebenden Systemen und essentiell für alle biologischen Prozesse. Hierzu gehören Stofftransport, mechanische Bewegung, Übermittlung von Nervenimpulsen, Immunsystem, Kontrolle des Wachstums sowie Differenzierung von Zellen. Proteine sind lineare Polymere aufgebaut aus Aminosäuren. Sie verkörpern den Übergang der eindimensionalen Welt ihrer Aminosäuresequenz in die dreidimensionale Welt der Makromoleküle. Die verschiedenen Aminosäuren liefern eine große Vielfalt an funktionellen Gruppen wie Alkohole, Thiole, Thioethern, Carbonsäuren, Carboxamiden wie auch eine Vielfalt von basischen Gruppen. Die chemische Reaktivität der Proteine basiert auf der Anordnung dieser funktionellen Gruppen und ist essentiell für die Funktion als Enzym[141].

Jede α-Aminosäure besteht aus einem zentralen α-Kohlenstoffatom (C_α), welches mit einer Carbonsäuregruppe (COOH), einer Aminogruppe (NH_2), einem Wasserstoffatom (H) und einer spezifischen Restgruppe (R) kovalent verknüpft ist. Aufgrund der Verknüpfung des α-Kohlenstoffatoms mit vier verschiedenen chemischen Gruppen sind alle α-Aminosäuren außer Glycin chiral. Nur die L-Form der 20 proteinogenen α-Aminosäuren bildet einen Bestandteil von Proteinen[141]. Abbildung 2.10 zeigt alle 20 proteinogenen Aminosäuren nach physikalisch-chemischen

Eigenschaften gruppiert.

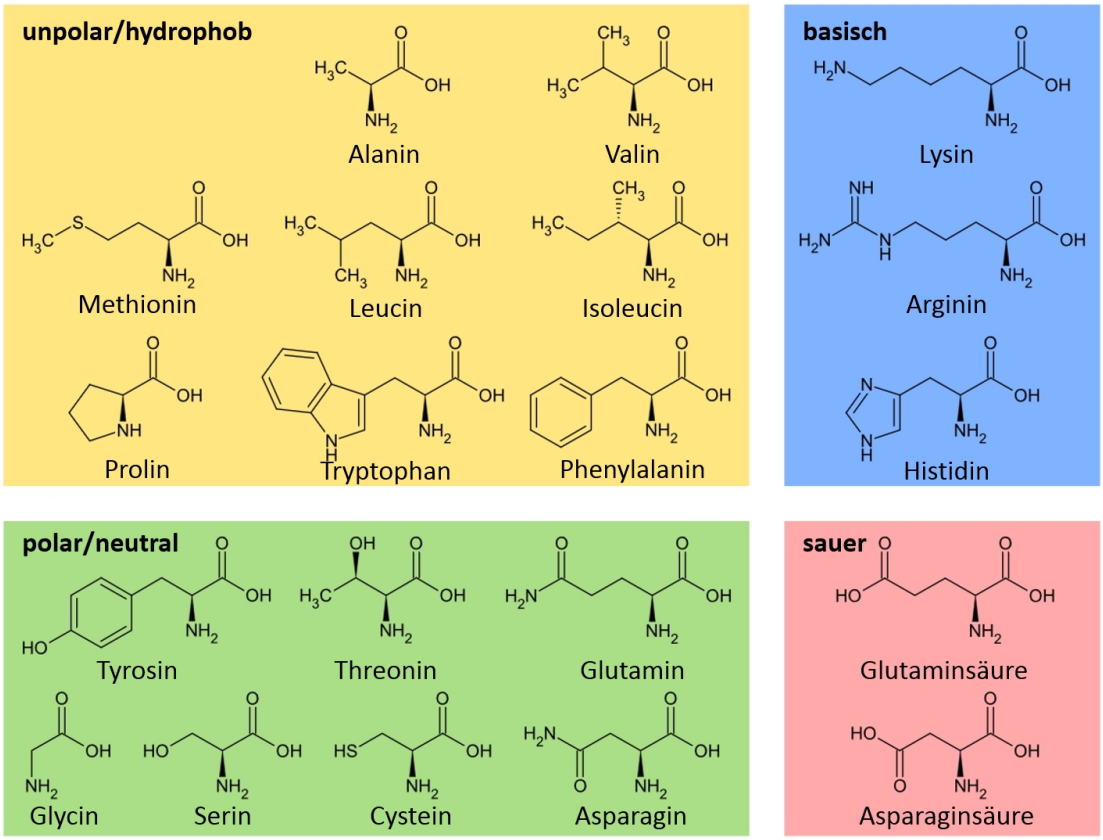


Abbildung 2.10: *Darstellung aller 20 proteinogenen Aminosäuren nach physikalisch-chemischen Eigenschaften gruppiert*[142].

Die räumliche Organisation wird in vier Ebenen eingeteilt. Die erste Ebene ist die Primärstruktur, bestimmt durch die spezifische Aminosäuresequenz der Polypeptidkette des Proteins.

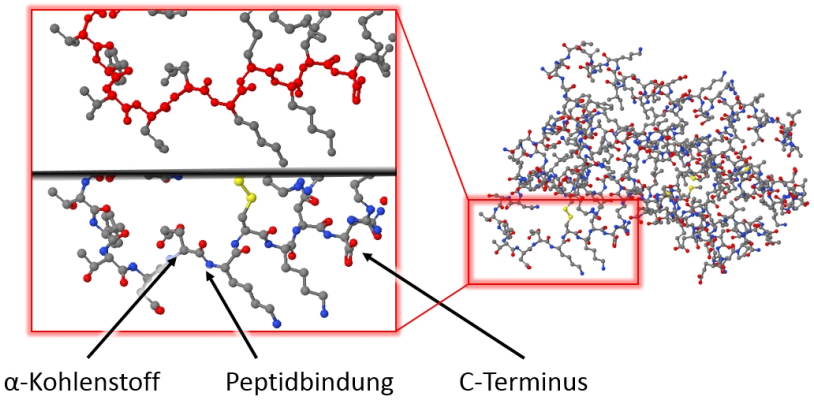


Abbildung 2.11: *Darstellung der Primärstruktur eines Proteins. Links oben sind die Atome des Proteinrückgrats rot hervorgehoben.*

Die nächsthöhere Organisationsebene ist die Sekundärstruktur, welche sich aus häufig auftretenden Strukturmerkmalen zusammensetzt. Dazu gehören die α-Helix, das β-Faltblatt sowie die β-Kehre. Diese Strukturen bilden und stabilisieren sich über Wasserstoffbrücken zwischen den Peptidbindungen. Die dritte Stufe ist die Tertiärstruktur. Sie beschreibt die finale Faltung des Proteins. Neben den nicht-bindenden Kräften (Coulomb-Wechselwirkungen, Van-der-Waals-Kräfte und Wasserstoffbrücken) zur Stabilisierung der Proteinkonformation findet man auch kovalente Disulfidbrücken zwischen Cysteinresten. Proteine bestehen häufig nicht nur aus einer monomeren Einheit, sondern aus agglomerierten Proteinkomplexen. Diese Zusammenlagerung von Proteinuntereinheiten nennt man Quartärstruktur. Abbildung 2.12 zeigt alle Organisationsebenen anhand eines Insulindimeres (PDB ID: 4INS)[143].

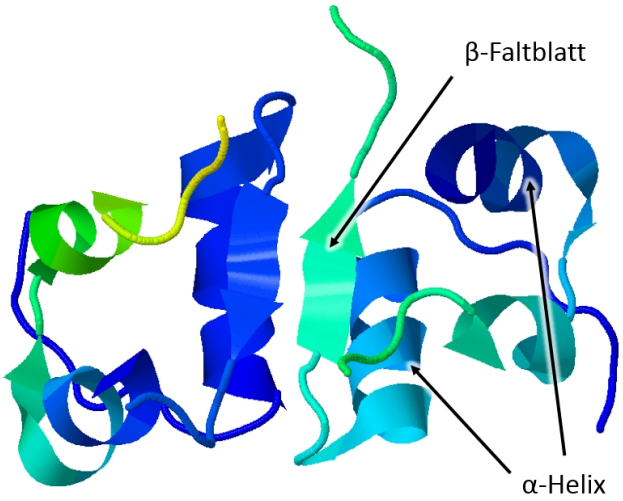


Abbildung 2.12: *Darstellung der Sekundär-, Tertiär- und Quartärstruktur anhand eines Insulindimeres.*

Im Gegensatz zur Peptidbindung sind die Bindungen der Aminogruppe (N) bzw. der Carboxylgruppe (C') zum α-Kohlenstoffatom (C_α) reine Einfachbindungen. Um diese beiden Bindungen können die benachbarten Peptidbindungen rotieren. Der Rotationswinkel um die N-C_α-Achse wird als ϕ und der Rotationswinkel um die C_α-C'-Achse als ψ gekennzeichnet. Nicht alle ϕ und ψ sind aufgrund der sterischen Blockierung durch die Nachbaratome erlaubt. Abbildung 2.13 zeigt die erlaubten Winkel von ϕ und ψ in einem Ramachandran-Diagramm. Die anhand durchgezogener Linien hervorgehobenen Winkel zeigen die favorisierten Regionen in Proteinen, gefolgt von den erlaubten Winkeln, dargestellt durch gestrichelte Linien. In den Flächen außerhalb dieser gekennzeichneten Regionen kommen sich die Atome näher als ihr Van-der-Waals-Radius. Diese Konformationen sind folglich energetisch ungünstig.

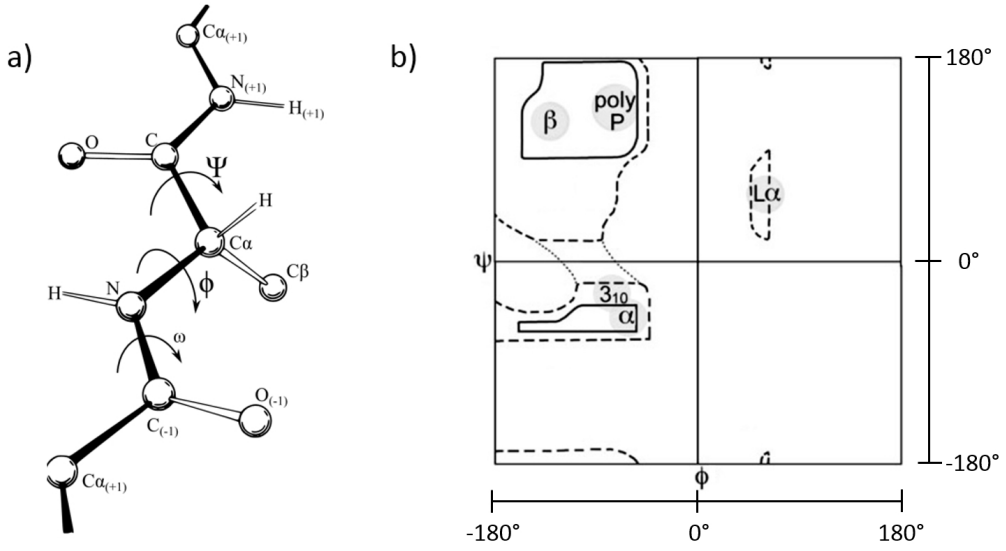


Abbildung 2.13: *a) Proteinrückgrat Diederwinkel ϕ und ψ. b) Ramachandran-Diagramm*[144].

Bemerkenswert an dieser Stelle ist, dass die Sekundärstrukturmotive β-Faltblatt und α-Helix ausschließlich in den favorisierten Regionen anzutreffen sind und somit energetisch günstige Konformationen darstellen. Die Winkel der β-Faltblatt Konformation sind $\phi \approx -139$ und $\psi \approx -135$, die der α-Helix Konformation sind $\phi \approx -57$ und $\psi \approx -47$. Die Einschränkung der Winkel von ϕ und ψ minimiert hinreichend die Zahl der möglichen Strukturen, sodass die Proteinfaltung überhaupt erst möglich wird[141].

2.4 Proteinaufreinigung mit Zinkricinoleat

Zinkricinolat (CAS-Nummer: 13040-19-2; IUPAC-Notation: Zink(Z,12R)-12-hydroxyoktadek-9-enoat) gehört zu der Substanzklasse der metallischen Seifen[72,145]. Zinkricinoleat repräsentiert somit ein Salz einer Fettsäure, oder präziser, das Zinksalz der Ricinolsäure. Es wird im weiteren Verlauf als $Zn(Ri)_2$ abgekürzt.

Die chemische Struktur besteht aus einem Zn^{2+}-Ion, flankiert von zwei Molekülen Ricinolsäure (siehe Abbildung 2.14). Die Ricinolsäure besitzt eine Doppelbindung in cis-Konfiguration zwischen dem 9. und 10. Kohlenstoffatom. Eine

Besonderheit in der Struktur ist eine Hydroxylgruppe am 12. Kohlenstoffatom. Keine andere natürlich vorkommende Fettsäure besitzt eine derartige Hydroxylgruppe[146,147].

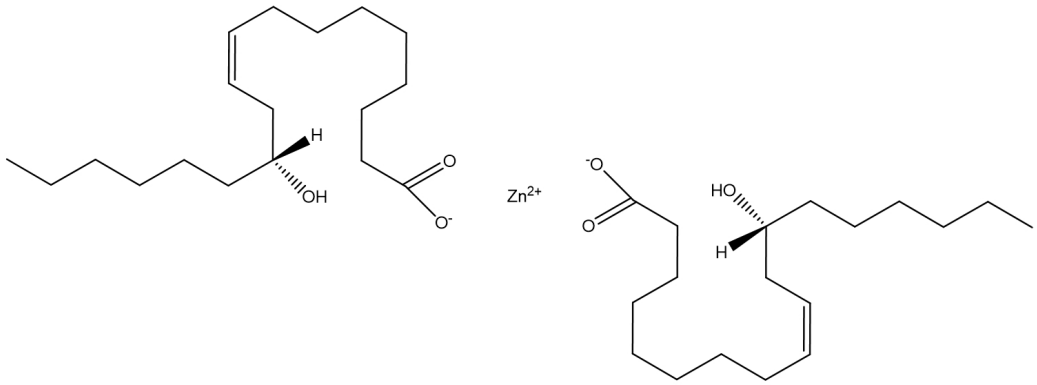


Abbildung 2.14: *Chemische Struktur des Zinkricinoleats.*

Aufgrund der geringen Toxizität ($LD50_{oral} > 2000$ mg·kg^{-1}) und der fehlenden Sensibilisierung von Haut und Augen ist Zinkricinoleat perfekt für die Anwendung im Haushalt, in der Kosmetik und der pharmazeutischen Industrie[68–72]. Eine hohe biologische Abbaurate von 99 % und das Fehlen jeglicher antibakteriellen sowie fungiziden Wirkung[68,72] trägt zusätzlich zur Verwendbarkeit in den oben genannten Bereichen bei.

Die Aufreinigung von Proteinen kann mittels verschiedenster Methoden wie Ionenaustauschchromatographie oder hydrophober Interaktionschromatographie durchgeführt werden[148–150]. Eine spezielle Variante für die Proteinaufreinigung ist die Metallchelatchromatographie (*Immobilized metal ion affinity chromatography*, kurz IMAC). Der Aufbau der Matrix der IMAC-Säule besteht aus einem festen Absorptionsmittel (Silikagel), welches mit einem Komplexierungsagens für Metallionen, z.B. mit Derivaten der Iminodiessigsäure, funktionalisiert wird. Anschließend werden die entsprechenden Metallionen zu dem funktionalisiertem Medium gegeben[151].

Die Affinität zwischen dem Analyten und der Säulenmatrix entsteht durch elektrostatische Wechselwirkungen zwischen Metallionen und beispiels-

weise Polyhistidin-Tag-Proteinen. Diese Proteine oder Enzyme sind genetisch derart modifiziert, sodass sie N- oder C-terminal mehrere Histidinmoleküle tragen[69,152–155]. Die so veränderten Proteine nennt man kurz His-Tag-Proteine. Abbildung 2.15 zeigt eine Vergrößerung des His-Tags eines Fibroblasten-Wachstumsfaktorproteins (FGF-1, PDB-ID: 3UD7[156]) und die molekulare Struktur der Aminosäure Histidin.

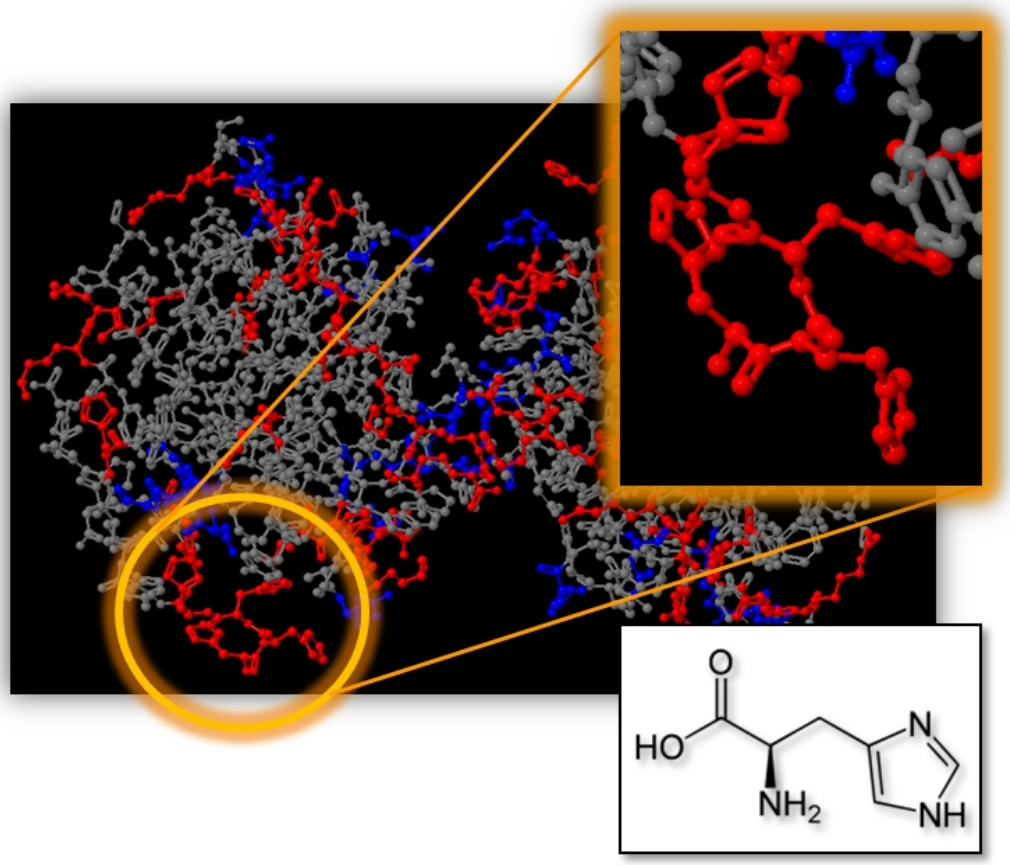


Abbildung 2.15: *Vergrößerte Darstellung eines terminalen His-Tags. Unten rechts ist die molekulare Struktur der Aminosäure Histidin abgebildet.*

Für die Aufreinigung von His-Tag-Proteinen mit der IMAC-Methode muss die Säulenmatrix mit Nickelionen funktionalisiert sein. Wenn ein His-Tag-Protein auf ein gebundenes Nickelion trifft, wird dieses durch einen Komplexierungsprozess der Imidazolgruppe der Histidinmoleküle des Polyhistidin-Tags festgehalten. Die Elution der so festgehaltenen Proteine geschieht durch Zugabe an Überschuss eines konkurrierenden Moleküls - etwa Imidazol - in die Pufferlösung[151,157,158]. Abbil-

dung 2.16 zeigt den Prozess der Aufreinigung mittels IMAC.

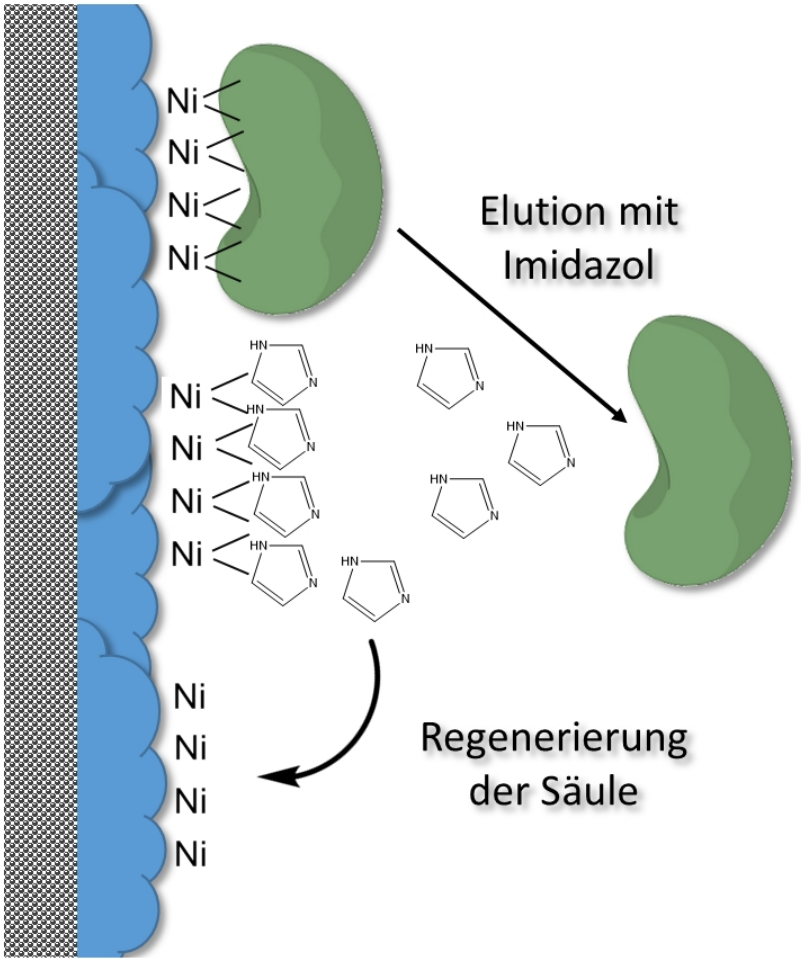


Abbildung 2.16: *Prinzip der Metallchelatchromatographie.*

Das Patent DE102011101880 A1[69] zeigt, dass $Zn(Ri)_2$ fähig ist, als neues Medium für die Aufreinigung von His-Tag Proteinen zu dienen. Es wird dargestellt, dass ein Substrat von Silikagel oder Cellulose mit $Zn(Ri)_2$ funktionalisiert werden kann. Mittels der so funktionalisierten Matrix konnte eine Aufreinigung des Proteins *Bacillus stearothermophilus* Esterase (Bste), welches ein His-Tag C-terminal trägt, durchgeführt werden. Zusätzlich ist die Funktionalisierung auch auf Oberflächen aus Aluminiumoxid durchführbar[69].

Aufgrund der geringen Toxizität und der biologischen Abbaubarkeit ist $Zn(Ri)_2$ besonders für die Lebensmittelproduktion geeignet. Toxische Schwermetallionen

wie z.B. Ni^{2+} oder Co^{2+}, die Verwendung in anderen Aufreinigungsprozessen (z.B. IMAC) finden, können durch weniger toxische Metallionen wie z.B. $Fe^{3+,\ 2+}$ oder Zn^{2+} ersetzt werden[69]. Nichtsdestotrotz ist der Mechanismus der Bindung von Polyhistidin-Tag-Proteinen an $Zn(Ri)_2$ bislang unbekannt.

2.5 Biologische Aktivität des Cyclotids Kalata B1

Cyclotide sind kleine Peptide die in Pflanzen vorkommen und eine Länge von 28-37 Aminosäuren haben[94,159–161]. Alle Cyclotide bestehen aus einem zyklischen Proteinrückgrat sowie drei konservierten Disulfidbrückenbindungen die eine Knotenstruktur bilden. Die einzelnen Disulfidbrücken bilden die Cysteinpaaren CI/CIV, CII/CV und CIII/CVI. Die Abschnitte zwischen den konservierten Cysteinen werden als Schlaufe 1 bis 6 bezeichnet. Dieses Strukturmotiv der Cyclotide wird auch *Cyclic cystine knot* (deutsch zyklischer-Cystein-Knoten, kurz CCK) genannt. Abbildung 2.17 skizziert die molekulare Struktur des Kalata B1 Cyclotids, welches als Prototyp dieser Art von Peptiden angesehen wird[89].

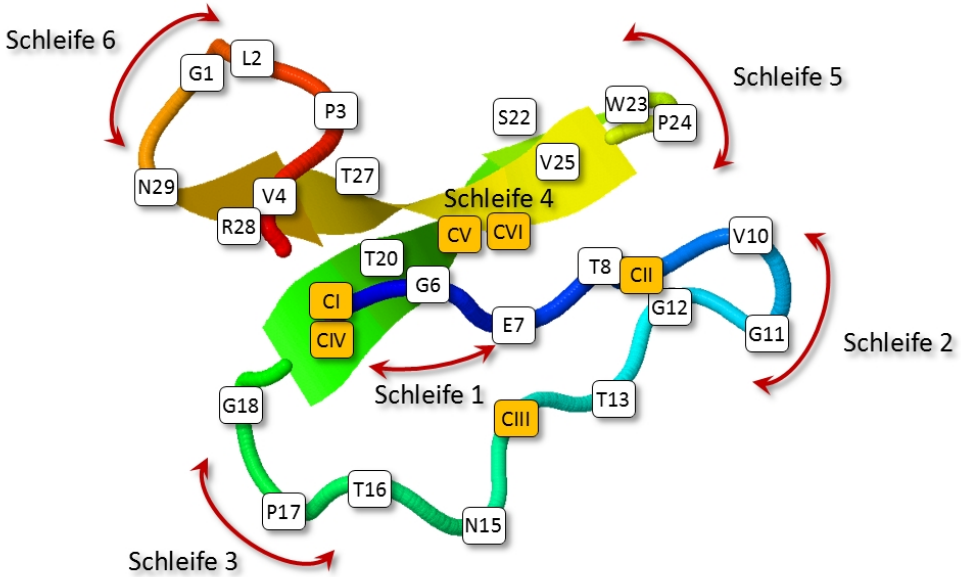


Abbildung 2.17: *Aufbau des Cyclotids Kalata B1 in der Cartoon-Darstellung. Zu sehen ist die Nummerierung der 29 Aminosäuren sowie die sechs Schlaufen. Die sechs konservierten Cysteine (orange) der Disulfidbrücken sind mit römischen Zahlen gekennzeichnet.*

Cyclotide zeigen einen großen Umfang an biologischen Aktivitäten. Diese umfassen: Anti-HIV, Neurotensin Antagonist, hämolytisch, antimikrobiell, Fäulnis verhütend und eine pestizide Wirkung. Aufgrund dieses Wirkungsspektrums werden Cyclotide als Defensine eingestuft[92].

Die Zytotoxizität der Cyclotide entsteht durch ihre membranzerstörenden Eigenschaften. Die hohe Affinität zu Lipidmembranen resultiert aufgrund zweier Domänen innerhalb der Proteinstruktur, die den Kontakt zu der Membran vermitteln. Der sogenannte *hydrophobic Patch* (deutsch hydrophober Fleck) besteht aus den Aminosäuren L_2, P_3 und V_4 von Schlaufe 6 sowie W_{23}, P_{24} und V_{25} von Schlaufe 5 (siehe Abbildung 2.18, grün)[89–94]. Den zweiten Kontaktpunkt formen die beiden ionischen Aminosäuren G_6 in Schlaufe 1 und N_{29} in Schlaufe 6 (siehe Abbildung 2.18, rot und blau), die mit den geladenen Kopfgruppen der Phospholipide interagieren. Eine besonders starke Affinität von Kalata B1 besteht zu Membranen mit einem hohen Phosphatidylethanolamin (PE) Anteil, wobei gilt: Je mehr PE-Moleküle in der Membran vorhanden sind, desto größer ist die Affinität. Das Eindringen des Moleküls in die Membran vermittelt die elektrostatische Wechselwirkung zwischen den beiden ionischen Aminosäuren und der PE-Kopfgruppe. G_6 steht dabei im Kontakt mit der positiv geladenen Ammoniumgruppe. N_{29} bildet den Kontakt mit der negativ geladenen Phosphatgruppe des Phosphatidylethanolamins aus[89,91,93,162].

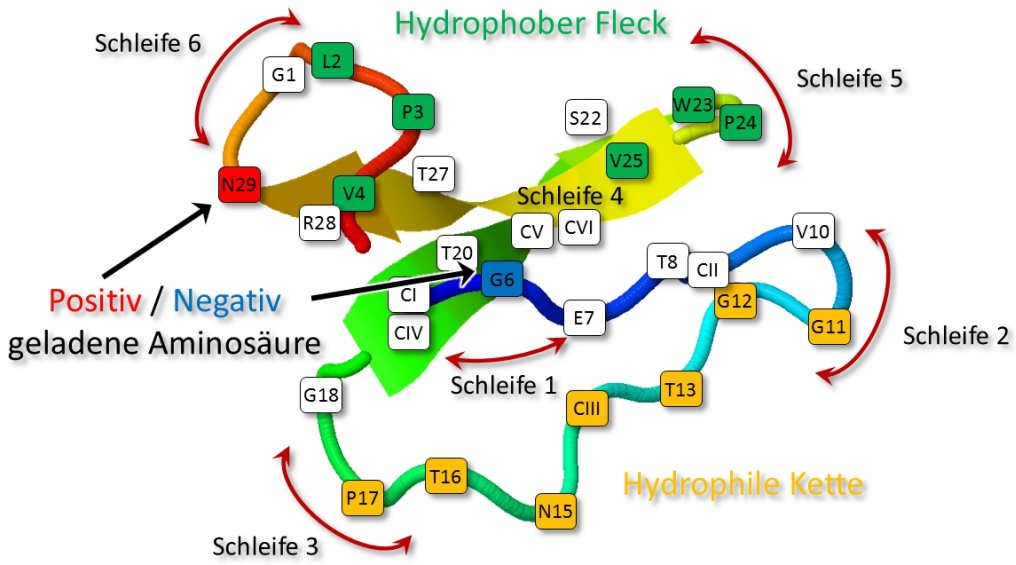


Abbildung 2.18: *Darstellung der einzelnen Domänen des Cyclotids Kalata B1.*

Kalata B1 trifft entweder als Monomer oder Oligomer auf die Membran. Die Wechselwirkung mit der Membran findet über den ausgeprägten *hydrophobic Patch* und den geladenen Residuen statt. Es wird davon ausgegangen, dass eine höhere Konzentration an Kalata B1 Molekülen die Eindringtiefe der Monomere erhöht. Die gebundenen Kalata B1 Moleküle diffundieren dann lateral über die Membran und oligomerisieren zu größeren Einheiten. Ab einem gewissen Oligomerisationsgrad tritt anschließend die Porenbildung ein, wobei angenommen wird, dass die Pore aus einem Ring mit 4 - 6 Tetrameren besteht. Der Durchmesser dieser Pore beträgt ungefähr 41 - 47 Å. Zwei Arten von Porenstrukturen werden postuliert: Die Fasslamellenpore und die Ringkernpore. Erstere entsteht hauptsächlich bei der Oligomerisation und zeigt eine fassartigen Aufbau. Die Struktur der Ringkernpore bildet sich durch eine Krümmung der Membran aus, bis hin zum Kontakt der Kopfgruppen der Phospholipide der gegenüberliegenden Seiten[89,92,95]. Des Weiteren kann beobachtet werden, dass Kalata B1 Moleküle bei höheren Konzentrationen Lipidmoleküle aus der Membran extrahieren und für eine Krümmung der Membran sorgen. Dies resultiert in einer verminderten Dicke der Doppelschicht. Die Solubilisierung der Membranlipide und die abnehmende Dicke der Lipiddoppelschicht führen letztendlich zum Bruch der Membran[91,96]. Abbildung 2.19 skizziert den Vorgang der Bindung, Oligomerisation, Porenbildung und anschließender Solubi-

lisierung der Phospholipide einer Biomembran durch Kalata B1.

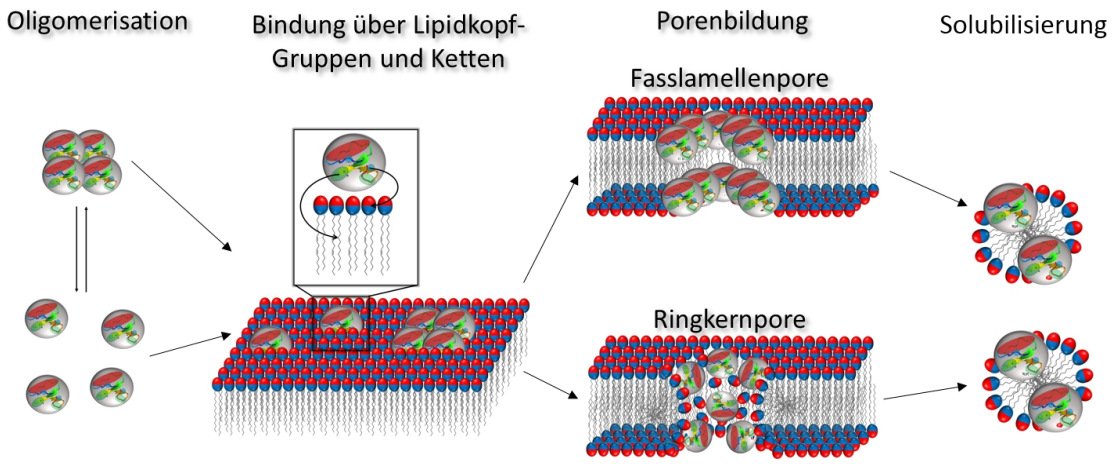


Abbildung 2.19: *Darstellung der Bindung, Oligomerisation, Porenbildung und anschließender Solubilisierung der Phospholipide einer Biomembran durch Kalata B1.*

3 Methodenentwicklung

In diesem Kapitel werden die entwickelten Methoden für die Simulation von Peptiden, Proteinen und Biomembran mit der molekularen Fragmentdynamik beschrieben. Der erste Teil beschreibt das verwendete Fragmentset inklusive des erarbeiteten Fragmentierungsschemas für Aminosäuren sowie aller verwendeter Moleküle. Außerdem wird die Berechnung des DPD-Repulsionsparameters anhand von MD-Simulationen mit dem Softwarepaket Discover 96.0 / 4.0.0 der Firma MSI[103] skizziert. Der zweite Abschnitt erläutert die zur Einführung von Biopolymeren in die molekulare Fragmentdynamik entwickelten Programme, die Behandlung der Daten eines Fragmentsets sowie die vorgenommenen Erweiterungen der Simulationssoftware MFD-FormulaOne[54].

3.1 Entwicklung des Fragmentsets für Biopolymere

In der molekularen Fragmentdynamik werden Moleküle explizit aus Molekülfragmenten aufgebaut. Für die Repräsentation von molekularen Strukturen und deren Simulation in der molekularen Fragmentdynamik müssen drei Problemstellungen gelöst werden: (1) Entwicklung und Implementierung einer Notation für die Repräsentation von Molekülen innerhalb von MFD-FormulaOne, (2) Ausarbeitung der Fragmenttopologie bzw. des Fragmentierungsschemas aller verwendeten Moleküle sowie (3) die Einführung einer Methode zur Berechnung des DPD-Repulsionsparameters.

3.1.1 Die *f*SMILES-Notation

Die in dieser Arbeit verwendete *fragment*-SMILES-Notation (kurz *f*SMILES) basiert auf den weit verbreiteten SMILES-Strukturcode (*Simplified Molecular-input Line-entry System*) für die Beschreibung der Topologie chemischer Strukturen[42,43]. Als Beispiel sei hier das Molekül Vanillin (Abbildung 3.1) aufgeführt. Die SMILES-Notation von Vanillin ist „O=Cc1ccc(O)c(OC)c1“. Die Atome der Moleküle werden für den organischen Atomsatz B, C, N, O, P, S, F, Cl, Br und I ohne Klammern kodiert, alle übrigen Elemente werden in eckige Klammern gesetzt (z.B.: [Au] für Gold). Kleingeschriebene Kohlenstoffatome deuten die Zugehörigkeit zu einem π-Elektronensystem an. Einfachbindungen werden einfachheitshalber vernachlässigt. Mehrfachbindungen werden in steigender Reihenfolge mit '=', '#' und '$' kodiert. Runde Klammern beschreiben Verzweigungen und Zahlen hinter den Atomsymbolen stehen für Verknüpfungen zur Darstellung von Ringsystemen. Eine detaillierte Beschreibung der SMILES-Notation liefert Weininger[43].

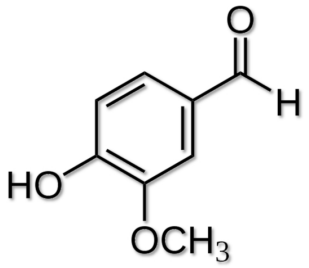


Abbildung 3.1: *Molekulare Struktur des Vanillins.*

Die *f*SMILES-Syntax beschreibt Moleküle entsprechend der SMILES-Notation in vereinfachter Form. Abbildung 3.2 zeigt das Fragmentierungsschema der Aminosäure Asparagin. Der zugehörige *f*SMILES lautet MeNH2BB(AcNH2)-HAcBB, wobei MetNH2BB für Methylamin, AcNH2 für Acetamid und HAcBB für Essigsäure steht. Fragmente, die das Proteinrückgrat aufbauen, werden mit einem BB-Suffix gekennzeichnet. In der *f*SMILES-Notation werden die einzelnen Atome der SMILES durch die jeweiligen Fragmentnamen ersetzt. Da Fragmentnamen keine einheitliche Länge aufweisen, müssen Bindungen zwischen den Fragmenten immer

mit einem Bindestrich ausgeschrieben werden. Um intramolekulare Verknüpfungen zu definieren, müssen die Indices der Verknüpfungen - im Gegensatz zum SMILES - mit eckigen Klammern umschlossen werden. Zahlen vor den Fragmentnamen stehen für die entsprechend häufige Aneinanderreihung dieser Fragmente. Die Beschreibung von Verzweigungen erfolgt auch bei den *f*SMILES mittels runder Klammern. Alle Regeln der *f*SMILES-Notation sind im Anhang 11.3 aufgestellt.

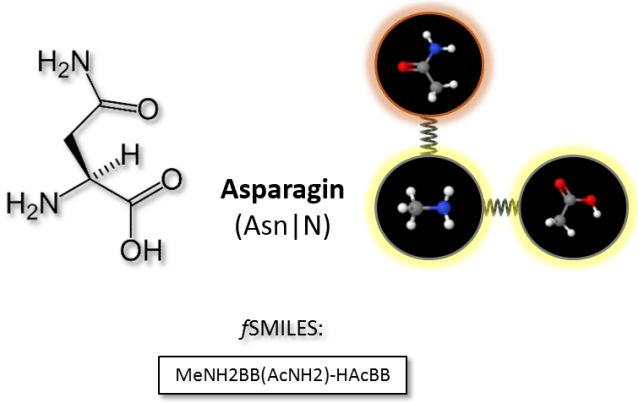


Abbildung 3.2: *Links: Molekulare Struktur; rechts: Fragmentierungsschema; unten: fSMILES der Aminosäure Asparagin.*

3.1.2 Fragmentierungsschema der verwendeten Moleküle

Zwei Anforderungen stellen sich an das hier zu entwickelnde Fragmentset: Der Aufbau aller Moleküle (Aminosäuren, Phospholipide und Zinkricinoleat) aus einer möglichst geringen Anzahl an Molekülfragmenten und eine chemisch sinnvolle Beschreibung der Moleküle anhand der verfügbaren Molekülfragmente. Während der Entwicklung des Sets hat sich gezeigt, dass ein Molekül nicht einfach in passende Molekülfragmente zerlegt werden kann. Der Aufbau von Alkylketten anhand von Methanfragmenten in entsprechender Anzahl zur Kettenlänge führt bei ansich polaren Molekülen zu einer zu repulsiven Charakterisierung gegenüber anderen polaren Fragmenten bzw. Molekülen. Desweiteren zeigen sich Moleküle mit einer großen Alkylkettenlänge aufgrund der hohen Fragmentanzahl als räumlich zu ausgedehnt gegenüber Aminosäuren ohne aliphatische Seitenkette. Die durchschnittliche Fragmentanzahl sollte ungefähr gleich für jede Aminosäure sein

bzw. der räumlichen Ausdehnung der jeweiligen Aminosäure entsprechen. Aliphatische Ketten sollten somit aus längerkettigen Fragmenten wie beispielsweise Ethan, Propan, usw. aufgebaut werden.

Abbildung 3.3 zeigt alle Molekülfragmente des entwickelten Fragmentsets. Die farbigen Fragmente sind für den Aufbau von Aminosäuren verantwortlich. Grüne Fragmente stellen hydrophobe und rote Fragmente hydrophile Seitenkettenfragmente dar. Die gelben Fragmente sind α-Kohlenstofffragmente des Proteinrückgrats. Die grauen Fragmente sind zusätzliche Fragmente zur Beschreibung von Biomembranen und Zinkricinoleat in dieser Arbeit.

Das komplette Aminosäurenfragmentierungsschema ist in Abbildung 3.4 abgebildet. Zu sehen ist das Fragmentierungsschema der Aminosäuresequenz aller Aminosäuren inklusive des zugehörigen fSMILES. Die Färbung der Fragmente entspricht derjenigen der einzelnen Molekülfragmente (Abbildung 3.3).

Die Fragmentierungsschemata der nicht-Aminosäuremoleküle sind in den jeweiligen Kapiteln des Ergebnisteils aufgeführt.

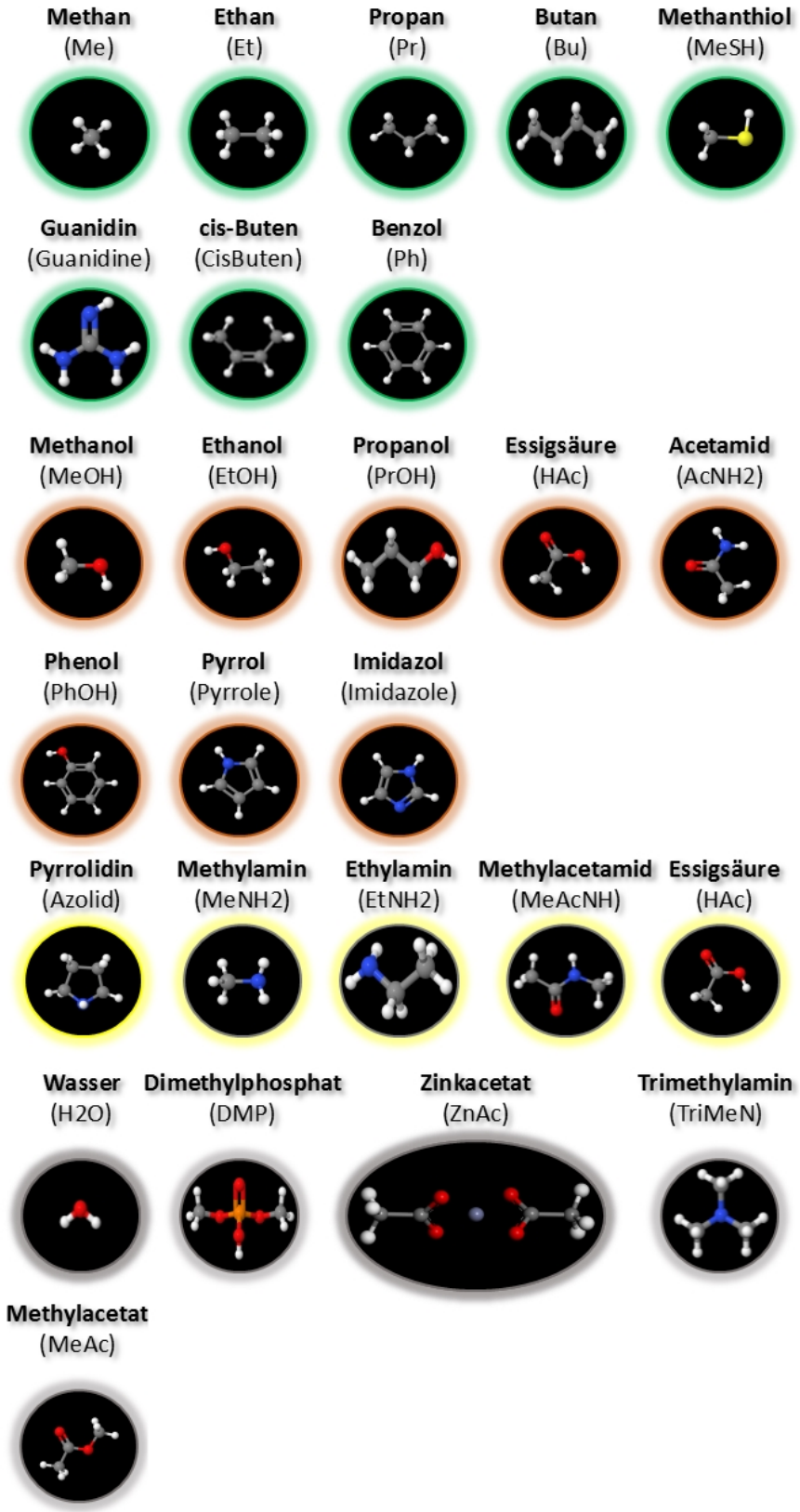
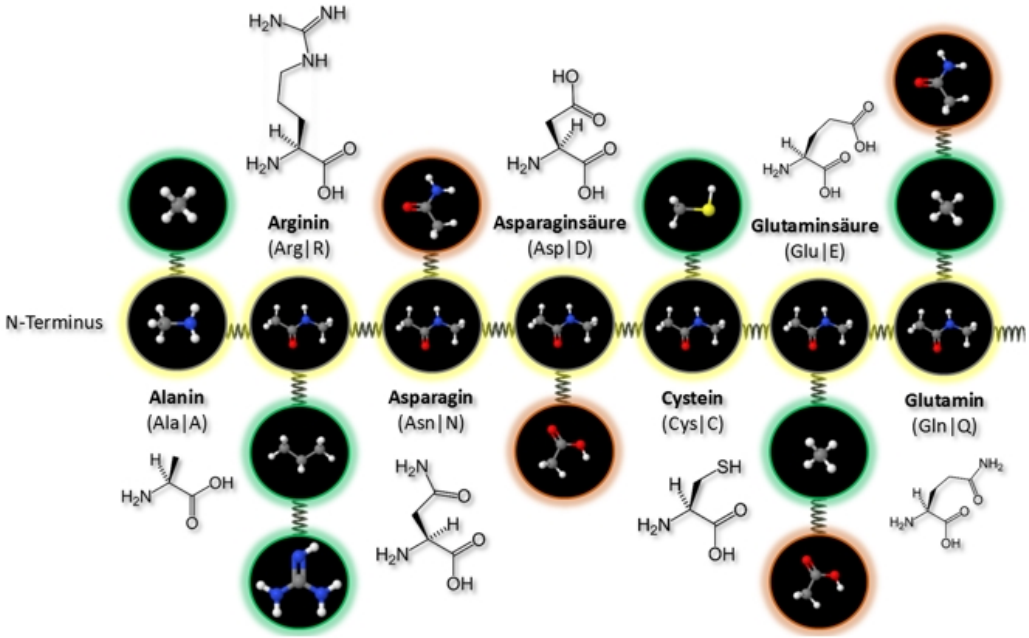
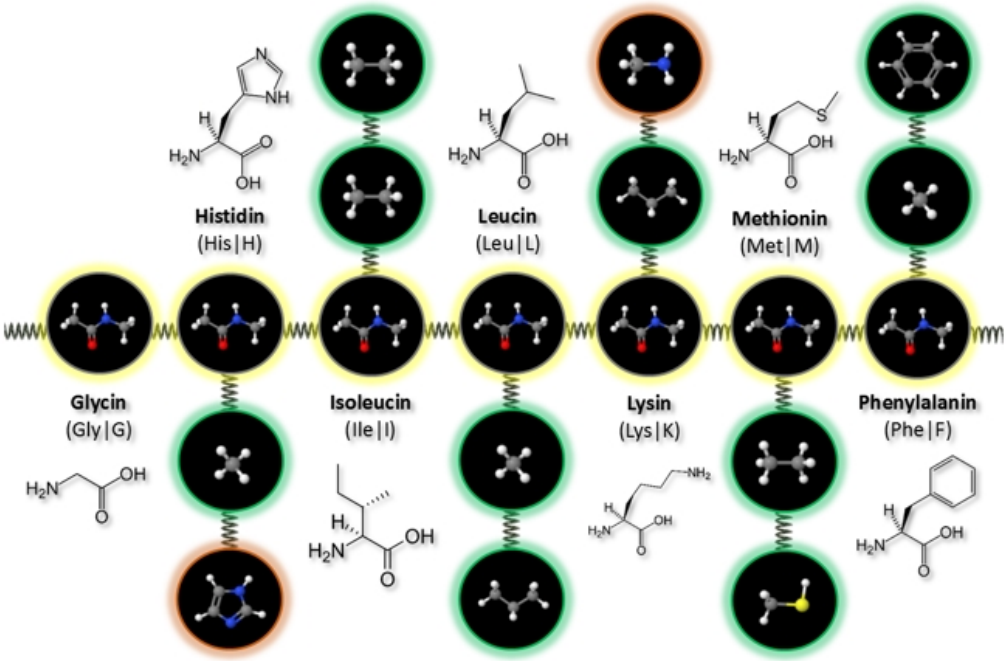


Abbildung 3.3: *Verwendete Molekülfragmente im Biopolymerfragmentset (grün: Hydrophobes Seitenkettenfragment, rot: Hydrophiles Seitenkettenfragment, gelb: α-Kohlenstofffragment, grau: Sonstige Fragmente). In Klammern steht die in fSMILES verwendete MFD-Abkürzung ohne explizites BB-Suffix.*



MeNH2PBB(Me)-**MeAcNH2BB(Pr-GuanidineP)**-MeAcNH2BB(AcNH2)-**MeAcNH2BB(HAcN)**-MeAcNH2BB (MeSH)-**MeAcNH2BB(Me-HAcN)**-MeAcNH2BB(Me-AcNH2)-HAcNBB



-MeNH2BB-**MeAcNH2BB(Me-Imidazole)**-MeAcNH2BB(Et-Et)-**MeAcNH2BB(Me-Pr)**-MeAcNH2BB(Et-MeSH)-**MeAcNH2BB(Me-Ph)-HAcBB**

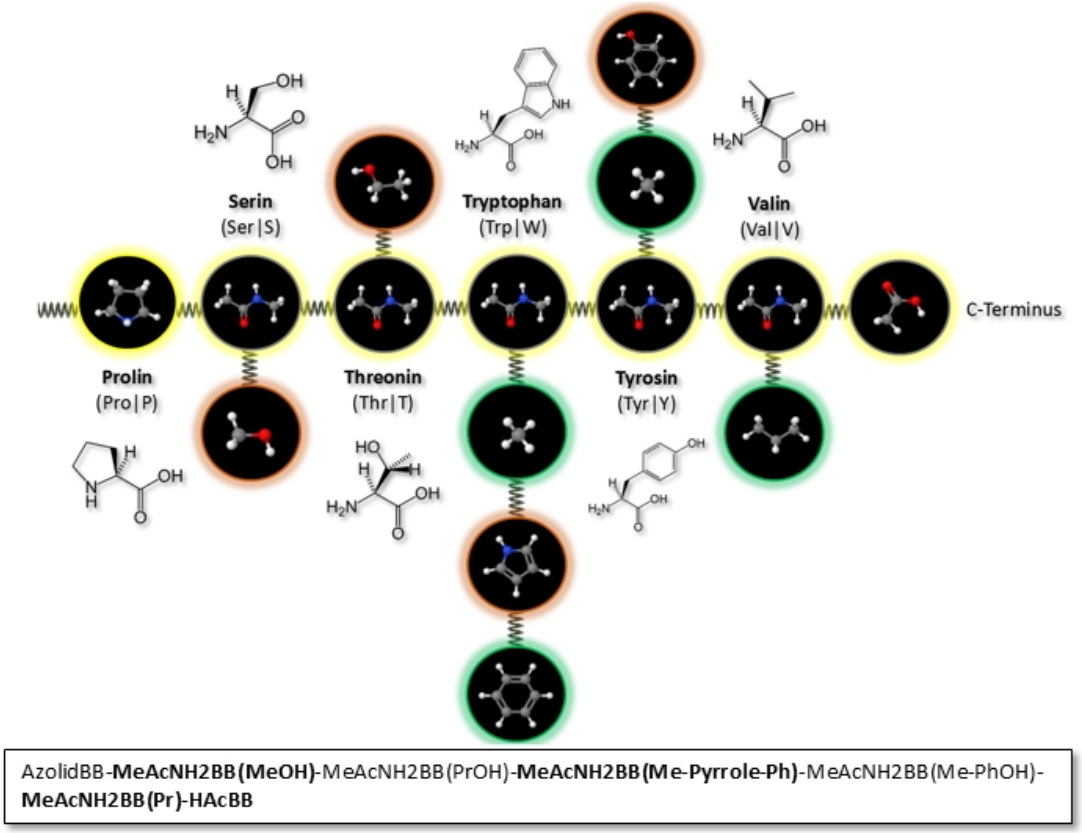


Abbildung 3.4: *Fragmentierungsschema aller 20 proteinogener Aminosäuren in Form einer Aminosäuresequenz. Unter jedem Schema ist der jeweilige fSMILES der Sequenz abgebildet.*

3.1.3 Berechnung der paarweisen Wechselwirkungsenergien E_{ij}

Die paarweisen Wechselwirkungsenergien E_{ij} für die Berechnung des DPD-Repulsionsparameters a_{ij} werden unter Verwendung der Molekulardynamik bestimmt. Die nötigen MD-Simulationen für die Berechnung der Wechselwirkungsparameter werden mit der Simulationssoftware Discover 96.0 / 4.0.0 der Firma MSI[103] durchgeführt. Die angewendeten Kraftfeldparameter sind die des COMPASS-Kraftfeldes[100].

Zwei verschiedene Ansätze (Dimere- und Boxmethode) mit jeweils zwei unterschiedlichen Versuchsaufbauten werden in dieser Arbeit zur Bestimmung des DPD-Repulsionsparamters entwickelt und untersucht. Folgende Simulationsbedingungen gelten für alle durchgeführten Berechnungen: Vor jeder MD-Simulation

wird eine Energieminimierung des zu untersuchenden Systems durchgeführt. Die Minimierung erfolgt in zwei Schritten: Der erste Minimierungsschritt basiert auf der *Steepest-Descent*-Methode mit *convergence* = 1000.0. Der zweite nutzt die *Conjugate-Gradient*-Methode mit *convergence* = 10.0 und basiert auf Polak-Ribiére-Polyak[104,105]. Der Parameter für *line_search_precision* wird auf 0.1 gesetzt.

Die eigentliche MD-Simulation ist aus zwei Schritten aufgebaut. Der erste Schritt dient der Einstellung des thermodynamischen Gleichgewichts. Die Länge der ersten Dynamiksimulation beträgt 10^4 Femtosekunden mit Zeitschritten von 1.0 Femtosekunden. Die Geschwindigkeiten werden auf Basis der Maxwell-Boltzmann-Verteilung randomisiert. Die Messungen der paarweisen Wechselwirkungsenergien finden im zweiten Schritt der MD-Simulation statt. Die Länge der zweiten Dynamiksimulation beträgt $2{\cdot}10^4$ Femtosekunden mit Zeitschritten von 0.5 Femtosekunden. Der Pseudo-Atom-Dimeransatz läuft in der Gasphase ab, alle anderen Simulationen finden im thermodynamischen NVT-Ensemble statt. Das verwendete Integrationsschema basiert auf dem *Velocity-Verlet*-Algorithmus[115]. Die Temperatur des Systems liegt bei 310 K, wobei der verwendete Thermostat auf der *Velocity-Scaling*-Methode[108] basiert.

Für die Bestimmung der paarweisen Wechselwirkungsenergien E_{ij} werden die nicht-bindenden Wechselwirkungen (Coulomb-Wechselwirkung und Van-der-Waals-Wechselwirkungen) zwischen den Molekülen des betrachteten Molekülpaares in der Simulationsbox aufsummiert und anschließend der Mittelwert gebildet. Die Berechnung der Wechselwirkungsenergien E_{ij} erfolgt abschließend anhand von Gleichung 2.27.

Der Dimeransatz

Für den Dimeransatz sind zwei Methoden verfügbar. Die erste Methode (Pseudo-Atom-Dimeransatz) verwendet Moleküle mit jeweils einem Schweratom (nicht-Wasserstoffatom) als Pseudo-Atom gekennzeichnet. Die MD-Simulationen laufen in der Gasphase ab. Damit die Moleküle nicht auseinanderdriften, wird ein harmo-

nisches Potential mit einer bestimmten Distanz d auf die Pseudo-Atome der beiden betrachteten Moleküle gelegt. Die Distanz wird schrittweise von einer Minimaldistanz bis hin zu einer Maximaldistanz erhöht. Für jede Distanz wird eine eigene MD-Simulation für die Bestimmung der paarweisen Wechselwirkungsenergie durchgeführt. Abbildung 3.5 zeigt das Prinzip des Dimeransatzes mit Pseudo-Atomen (orange gekennzeichnet).

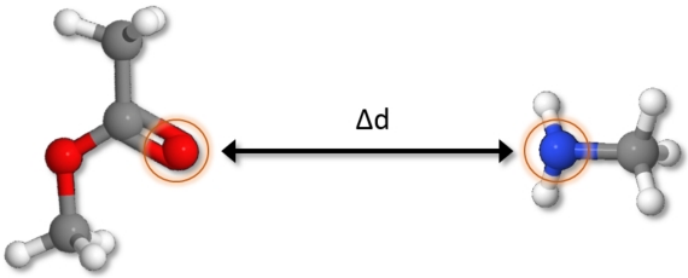


Abbildung 3.5: *Prinzip des Pseudo-Atom-Dimeransatzes. Eine mögliche Pseudo-Atom-Konfiguration ist orange hervorgehoben.*

Die Berechnung der paarweisen Wechselwirkungsenergie E_{ij} geschieht unter Verwendung folgender Formel:

$$E_{ij} = \frac{1}{N_{PS}} \sum_{n=1}^{N_{PS}} E_{ij,n} \tag{3.1}$$

wobei N_{PS} die Gesamtanzahl aller Pseudo-Atom-Konfigurationen und $E_{ij,n}$ die paarweise Wechselwirkungsenergie eines Molekülpaares mit einer spezifischen Pseudo-Atom-Konfiguration ist.

Der Rechenaufwand dieses Ansatzes steigt quadratisch im Bezug auf die Anzahl der eingesetzten Fragmente. In dem hier vorgestellten Fragmentset sind 23 Molekülfragmente enthalten. Nach der Bestimmung aller Pseudo-Atome jedes Moleküls erhöhte sich die Zahl der zu untersuchenden Moleküle auf 62. Symmetrien innerhalb der Moleküle werden berücksichtigt und nicht doppelt berechnet. Die Anzahl

der nötigen MD-Simulationen ist gegeben durch:

$$N = \frac{d \cdot (N_{MF}^2 + N_{MF})}{2 \cdot \Delta d} \tag{3.2}$$

wobei N_{MF} die Anzahl der Molekülfragmente, d die Distanz und Δd die Schrittweite ist. Betrachtet wird eine Distanz zwischen 1.0 Å und 10 Å mit einer Schrittweite von 0.05 Å, woraus sich $3.5 \cdot 10^5$ MD-Simulation für die Berechnung des Fragmentsets ergeben. Die Orientierung der Moleküle zueinander innerhalb der Simulation wird zufällig gewählt. Mit einer ungefähren Rechenzeit von 30 s pro Simulation kommt man auf eine Prozessorzeit von ungefähr 122 Tagen. Die Simulationen lassen sich zwar gut parallelisieren, jedoch skaliert die Aufteilung der einzelnen Rechnungen auf verschiedene Prozesse nicht linear mit der Anzahl der verwendeten Prozessorkerne.

Die Problematik dieses Ansatzes ist im Diagramm von Abbildung 3.6 dargestellt. Die paarweisen Wechselwirkungsenergien entsprechen den Energieminima der abgebildeten Verläufe. Die zu bestimmenden Energieminima sind häufig nur gering oder gar nicht ausgeprägt und unterliegen simulationsbedingten statistischen Schwankungen. Dies führt zu einer hohen Fehlerrate bezüglich der mit diesem Ansatz berechneten DPD-Repulsionsparameter und liefert kein konsistentes Fragmentset.

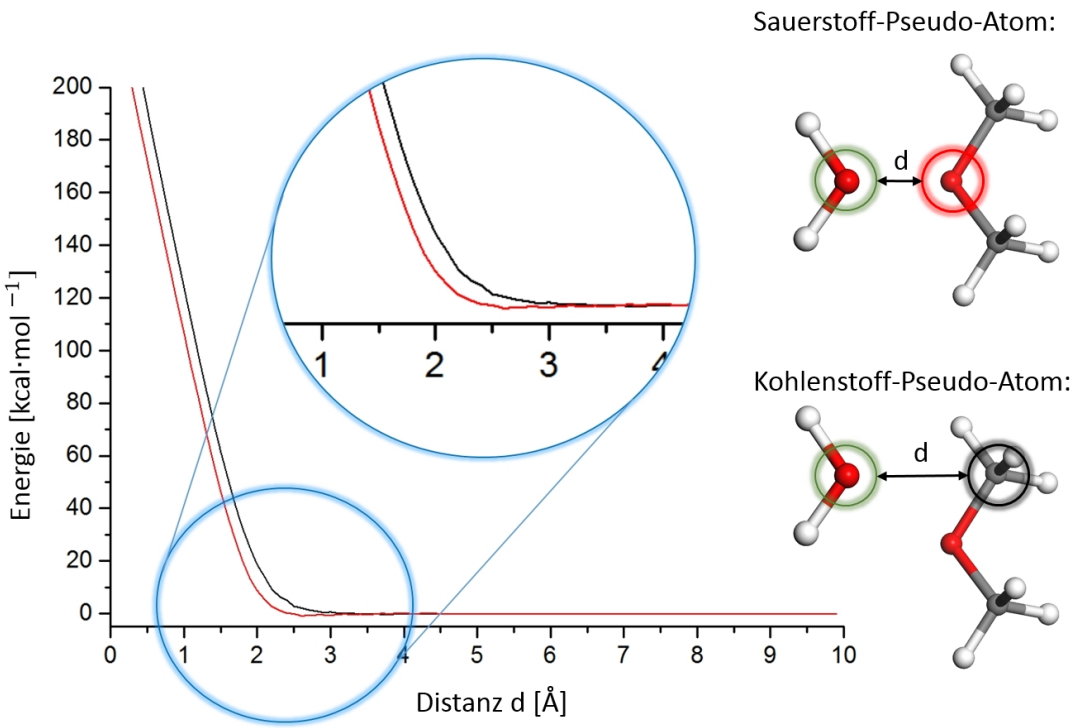


Abbildung 3.6: *Ergebnis einer Simulation des Molekülpaares Dimethylether-Wasser mit dem Pseudo-Atom-Dimeransatz. Rot ist das Ergebnis mit einem Sauerstoffatom bzw. schwarz mit einem Kohlenstoffatom als Pseudo-Atom im Dimethylether.*

Der zweite Dimeransatz funktioniert ähnlich zu dem vorherigen Ansatz. Hier wird nicht in der Gasphase simuliert, sondern eine Simulationsbox mit periodischen Randbedingungen verwendet. Die zu variierende Größe ist in diesem Fall die vorherrschende Dichte ρ der Simulationsbox. Die Dichte innerhalb der Simulationen wird über die Größe der Simulationsbox bestimmt, d.h. für eine höhere Dichte muss eine kleinere Kantenlänge der Box gewählt werden. Betrachtet wird eine Dichte von 0.05 g·ml^{-1} bis 3.0 g·ml^{-1} mit einer Schrittweite von 0.05 g·ml^{-1}. Abbildung 3.7 skizziert den Versuchsaufbau des dichteabhängigen Dimeransatzes.

Eine weitere Einstellung ermöglicht die Simulation ohne periodische Randbedingungen. Für die Bewerkstelligung wird ein harmonisches Potential auf die Grenzen der Box gelegt, wodurch die Moleküle im Zentrum der Simulationsbox gehalten werden.

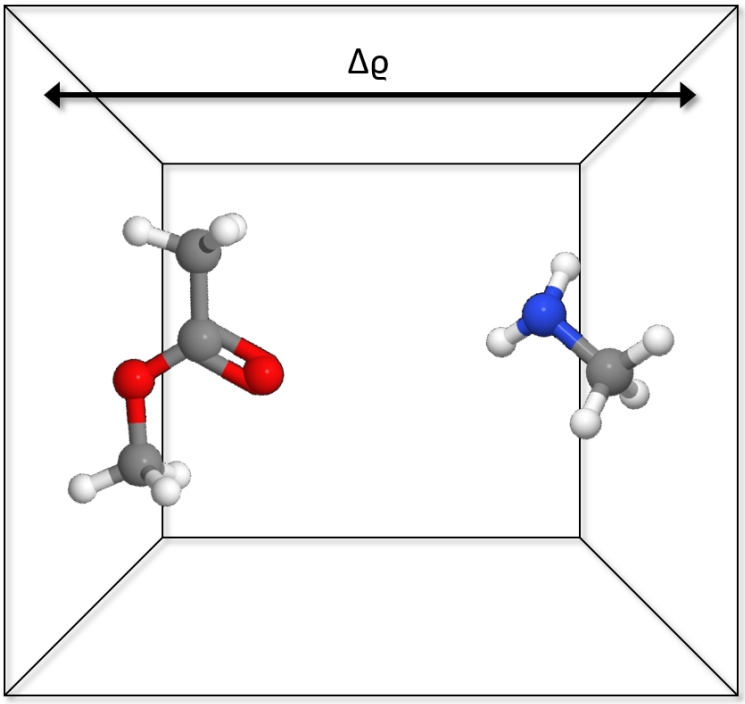


Abbildung 3.7: *Prinzip des dichteabhängigen Dimeransatzes.*

Bezüglich der Rechenzeit ist diese Methode dem Pseudo-Atom-Ansatz vorzuziehen, da nur Simulation in einem relativ geringen Dichtebereich (ungefähre Dichte von 0.75 bis 1.5 g/ml mit einer Schrittweite von 0.05 g/ml) durchgeführt werden müssen. Der betrachtete Dichtebereich sollte ungefähr der Dichte einer wässrigen Lösung entsprechen, damit die in dieser Arbeit auf Wasser basierenden Systeme realistisch abgebildet werden. Außerdem entspricht die Anzahl der betrachteten Molekülfragmente der ursprünglichen Anzahl von 23 Fragmenten. Die Anzahl der nötigen MD-Simulation errechnet sich äquivalent zur vorangegangenen Methode und ergibt einen Wert von 8280 Simulationen. Dies entspricht einer Prozessorzeit von ungefähr 3 Tagen bei einer Simulationszeit von 30 s pro MD-Simulation.

Das Problem an diesem Verfahren ist die Auswertung der Simulationsdaten bzw. die Bestimmung des Energieminimums. Abbildung 3.8 zeigt ein Diagramm der Molekülpaare Dimethylether-Wasser (rot) und Dimethylphosphat-Methylamin (schwarz). Für das Paar Dimethylether-Wasser ist ein Minimum leicht bestimmbar und es liegt in einem sinnvollen Dichtebereich (grün) für die Bestimmung des

DPD-Repulsionsparameters für wässrige Lösungen. Wohingegen das Molekülpaar Dimethylphosphat-Methylamin eine Vielzahl von Minima aufweist. Hierbei ist nicht eindeutig bestimmbar, welcher Energiewert zur Beschreibung der Wechselwirkung herangezogen werden muss. Die Auswertung ist aufgrund der Komplexität nicht ohne weiteres automatisiert durchführbar und die manuelle Auswertung von 8280 Simulationsergebnissen nicht in angemessener Zeit praktikabel.

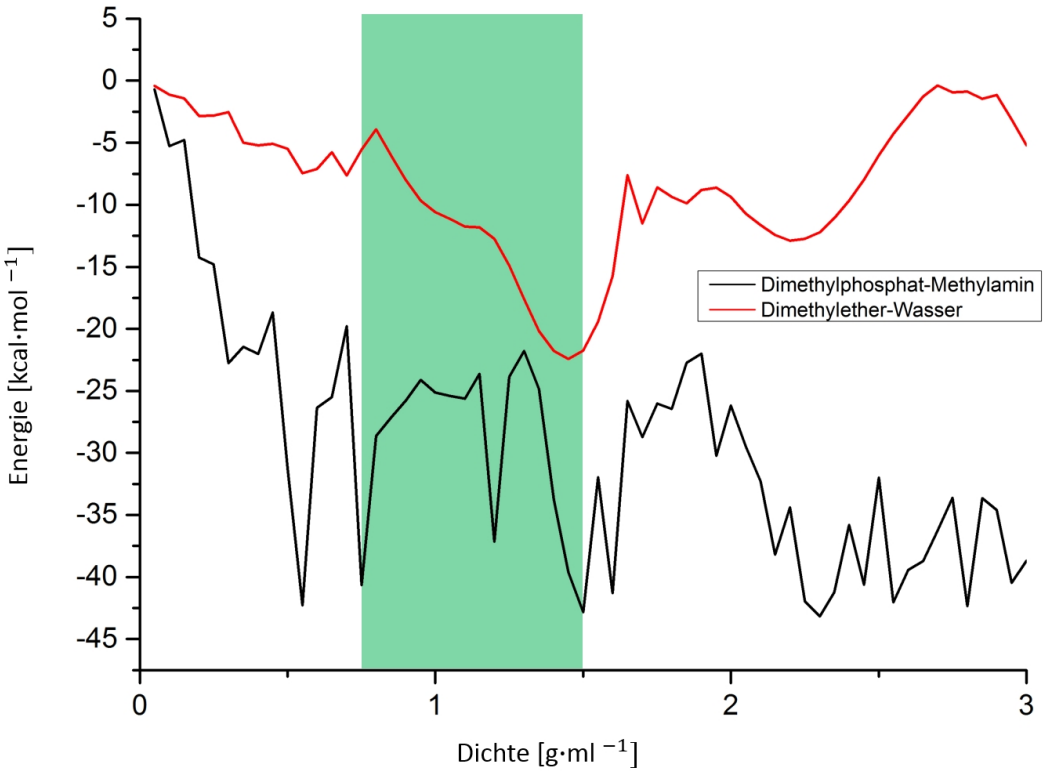


Abbildung 3.8: *Ergebnis einer Simulation des Molekülpaares Dimethylether-Wasser (rot) und Dimethylphosphat-Methylamin (schwarz) mit dem dichteabhängigen Dimeransatz. Grün zeigt den Bereich, der als realistisch für Simulation in wässriger Lösung angesehen werden kann.*

Auch hier können die periodischen Randbedingungen mittels harmonischer Potentiale auf den Simulationsboxgrenzen unterbunden werden. Ohne periodische Randbedingungen zeigen sich abermals die gleichen Probleme wie bei den Simulationen mit periodischen Randbedingungen.

Die Boxmethode

Die Untersuchung der Dimereansätze hat gezeigt, dass aufgrund der Schwierigkeiten in der Auswertung eine alternative Methode erforderlich ist. Die folgenden Ansätze untersuchen Systeme mit mehr als zwei Atomen. Zwei Simulationsboxkonfigurationen sollen betrachtet werden: (1) Ein Gemisch bestehend aus gleichen Anteilen der Moleküle des betrachteten Molekülpaares und (2) ein zentral platziertes Molekül, umgeben von Molekülen des jeweiligen Partnermoleküls.

Der erste Ansatz berechnet die paarweisen Wechselwirkungsenergien aus einem Gemisch mit gleichen Anteilen jedes Moleküls des betrachteten Molekülpaares. Abbildung 3.9 zeigt die Startgeometrie der Boxmethode mit einem Molekülgemisch bestehend aus H_2O (Wasser) und CH_3COOH (Essigsäure).

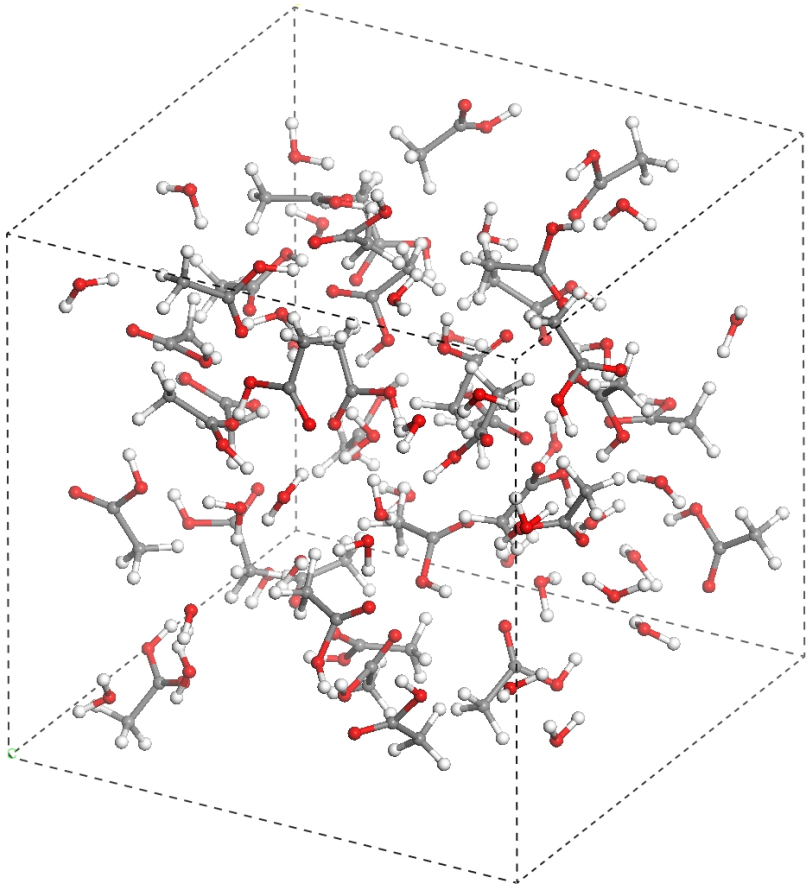


Abbildung 3.9: *Ansicht der MD-Simulationsbox für die Berechnung der paarweisen Wechselwirkungsenergie E_{ij} zwischen H_2O - CH_3COOH.*

Die Anzahl der nötigen MD-Simulationen errechnet sich nach der Formel:

$$N = \frac{N_{MF}^2 + N_{MF}}{2} \tag{3.3}$$

Im Gegensatz zu den Dimeremethoden muss nur eine Simulationsbox für die Berechnung der paarweisen Wechselwirkungsenergien betrachtet werden. Jedoch beträgt der Rechenaufwand pro Simulation nicht nur Sekunden, sondern liegt im Bereich von mehreren Minuten. Der Durchschnittswert der Rechenzeit liegt bei ungefähr 30 Minuten pro Versuch, woraus sich sich eine Prozessorzeit von ungefähr 6 d ergibt. Zusätzlich hat sich herausgestellt, dass die Energiewerte der Simulationen in Abhängigkeit von der Startgeometrie Schwankungen unterliegen. Der Grund hierfür ist, dass die Moleküle in einer zufälligen Anordnung in der Box platziert werden und es so zu unterschiedlichen Verteilungen bzw. zur Clusterbildung der einzelnen Molekülspezies kommt. Dies hat zur Folge, dass die Energiewerte gemittelt aus mehreren Simulationen zu bestimmen sind. Die Abbildungen 3.10 und 3.11 zeigen den Verlauf des kumulierten Mittelwertes der Energie der nichtbindenden Wechselwirkungen für die Systeme Wasser-Wasser, Wasser-Methanol, Wasser-Methan und Methanol-Methan. Der Nullpunkt der betrachteten Energien würde sich ergeben, wenn die Moleküle unendlich weit voneinander entfernt sind. Deutlich zu erkennen ist, dass für weniger als 25 Simulationen deutliche Abweichungen zum Mittelwert von 50 Simulationen erkennbar sind. Ab einer Simulationsanzahl von 25 wird der Verlauf des Graphen ruhiger und der kumulierte Mittelwert geht in den Mittelwert der Versuchsreihe über. Betrachtet man die Standardabweichugnen der paarweisen Wechselwirkungsenergien der Systeme, so ergeben sich für die rein hydrophilen Molekülspezies größere Abweichungen ($\sigma_{Wasser-Wasser} = 2.16$, $\sigma_{Wasser-Methanol} = 2.16$) als für Systeme mit mindestens einer hydrophoben Molekülspezies ($\sigma_{Wasser-Methan} = 0.50$, $\sigma_{Methanol-Methan} = 0.35$). Alle Molekülpaare mit der Möglichkeit zur wechselseitigen Ausbildung von Wasserstoffbrückenbindungen zeigen größere Energieabweichungen als Simulationen von Molekülpaaren ohne die Möglichkeit zur Ausbildung von Wasserstoffbrückenbindungen.

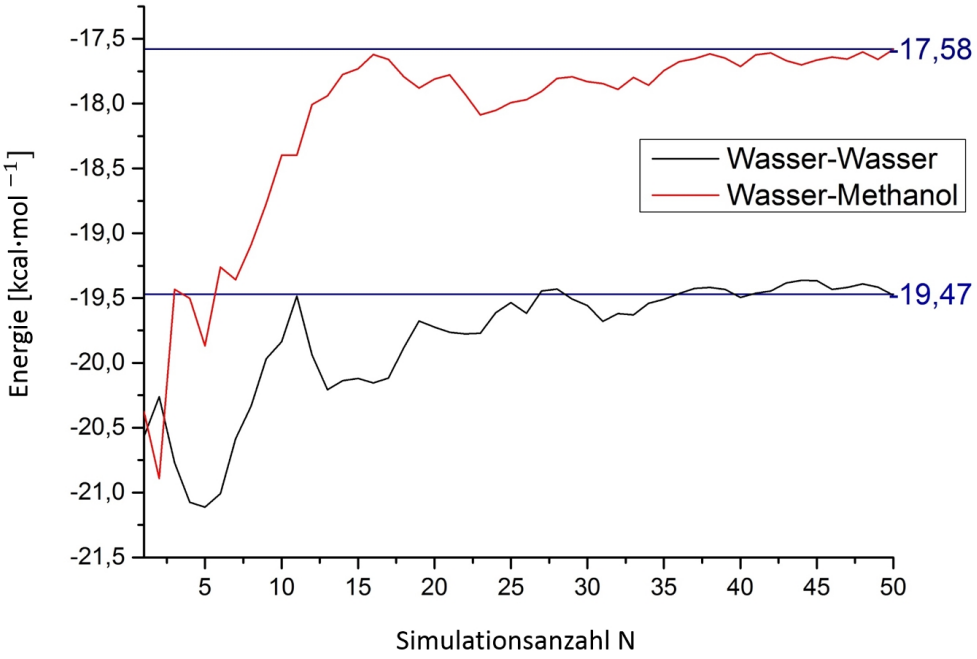


Abbildung 3.10: *Kumulierter Mittelwert der Systeme Wasser-Wasser und Wasser-Methanol.*

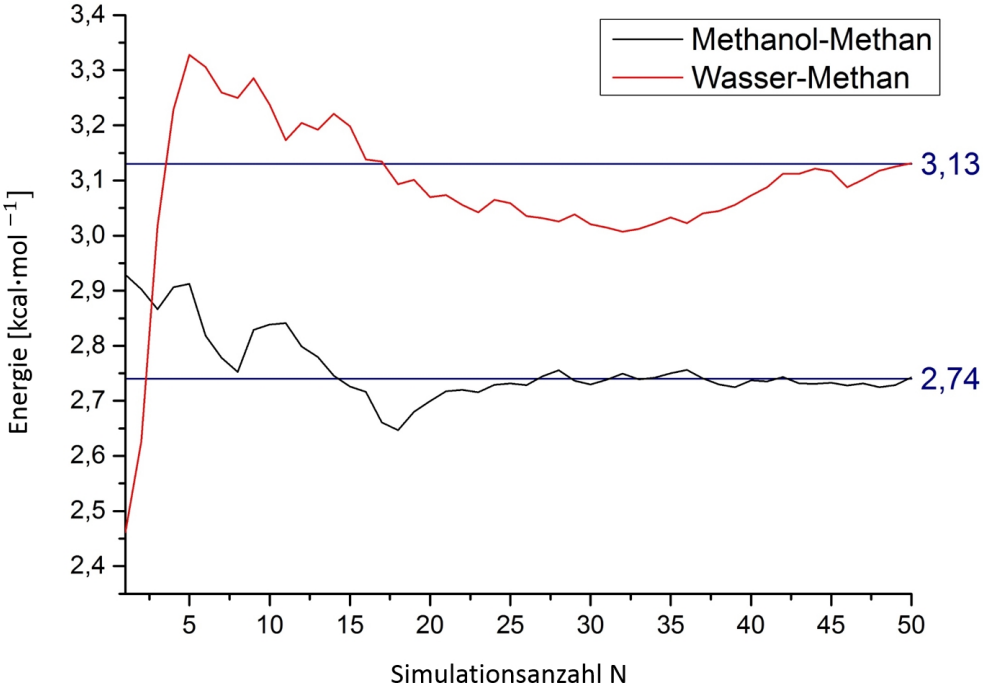


Abbildung 3.11: *Kumulierter Mittelwert der Systeme Wasser-Methan und Methanol-Methan.*

Setzt man die Wiederholungsrate der Simulationen auf einen Wert von 25 pro Molekülpaar, so erhöht sich die Simulationszeit auf 150 Prozessortage. Auch hier profitiert die Rechenzeit von der Parallelisierung auf mehrere Prozessorkerne, jedoch skaliert die Beschleunigung nicht linear mit der Anzahl der verfügbaren Kerne. Die berechneten DPD-Repulsionsparameter mit dieser Methode zeigen überwiegend erwartetes chemisches Verhalten. Einige Molekülgruppen zeigen hingegen physikalisch falsche Wechselwirkungen. Alkohole allgemein erwiesen sich beispielsweise als zu repulsiv zu polaren Fragmenten oder zu attraktiv zu unpolaren Molekülfragmenten. Dies führt zu einer falschen Beschreibung des chemischen Ensembles innerhalb von MFD-Simulationen, wodurch das berechnete Fragmentset nicht anwendbar ist.

Die zweite untersuchte Boxmethode verwendet eine Simulationsbox, bestehend aus insgesamt 50 Molekülen mit aktivierten periodischen Randbedingungen in allen drei Raumrichtungen. Für jedes Molekülpaar (Abbildung 3.12 zeigt Simulationsboxen für das System H_2O (Wasser) und CH_3COOH (Essigsäure)) sind jeweils zwei Rechnungen der homogenen Systeme (Abbildung 3.12a und 3.12b) sowie zwei Rechnungen der Mischsysteme (Abbildung 3.12c und 3.12d) durchzuführen. Die Boxen der Mischsysteme beinhalten zentral gelegen ein einzelnes Molekül der einen Spezies (blau hervorgehoben in Abbildung 3.12c und 3.12d) und als Lösungsmittel 49 Moleküle der anderen Molekülspezies. Die Platzierung der Moleküle erfolgt mit einem maximalen Abstand untereinander, d.h. die Moleküle werden ensprechend der hexagonalen dichtesten Kugelpackung angeordnet.

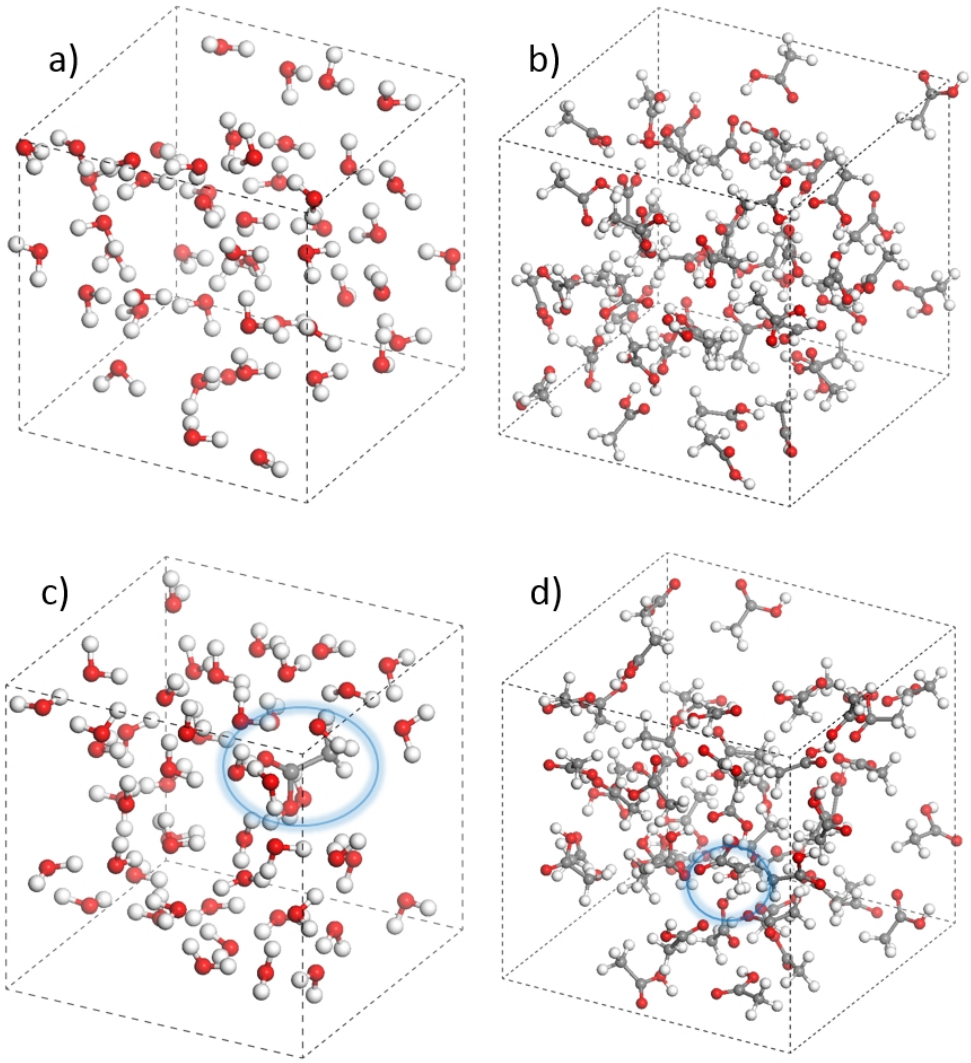


Abbildung 3.12: *Ansicht der MD-Simulationsboxen für die Berechnung der paarweisen Wechselwirkungsenergien E_{ij} (a: H_2O - H_2O, b: CH_3COOH - CH_3COOH, c: CH_3COOH - H_2O, d: H_2O - CH_3COOH; blau: Hervorhebung der jeweils einzelnen Molekülspezies).*

Im Gegensatz zu den vorher beschriebenen Methoden liegt bei den Mischsystemen keine Symmetrie vor, sodass $N = N_{MF}^2$ Simulationen durchgeführt werden müssen. Für 23 Molekülfragmente sind dies 529 MD-Simulationen. Die Dauer einer Simulation beträgt 30 Minuten, wobei jeder Versuch 25-mal wiederholt wird. Dies führt zu einer ungefähren Rechenzeit von 275 Prozessortagen. Die anhand dieses Verfahrens erhaltenen DPD-Repulsionsparameter zeigen alle das erwartete Verhalten bezüglich der attraktiven und repulsiven Wechselwirkung aller Molekülfragmente untereinander. Die anhand dieser Methode berechneten

Parameter werden in dieser Arbeit für die Simulation aller biologischer Systeme angewendet. Abbildung 3.13 zeigt die mit der Boxmethode berechneten DPD-Repulsionsparameter a_{ij}.

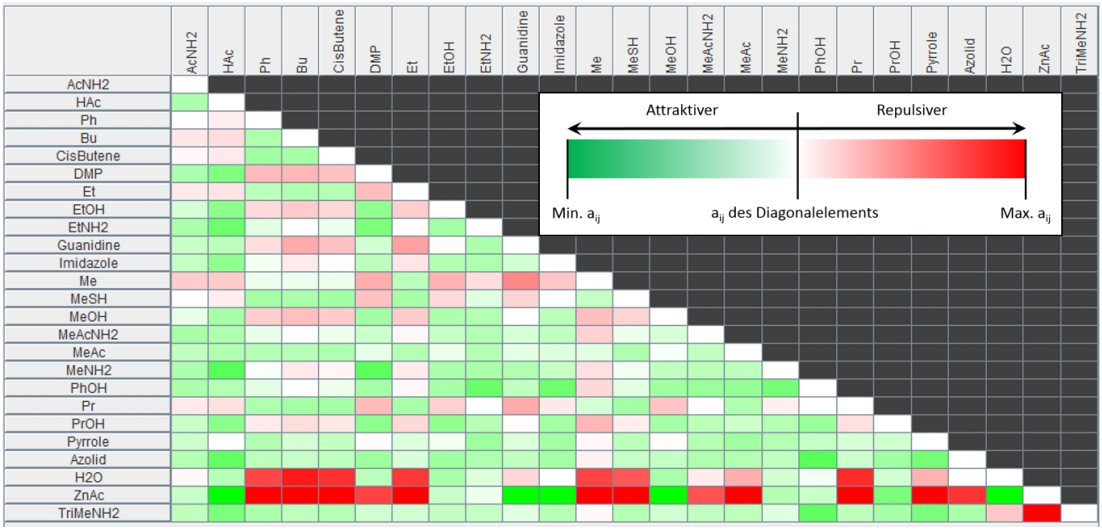


Abbildung 3.13: *Darstellung der DPD-Repulsionsparameter des MFD-Fragmentsets.*

3.1.4 Behandlung von geladenen Molekülfragmenten

Es hat sich gezeigt, dass die Berechnung von geladenen Fragmenten mit den im Kapitel 3.1.3 vorgestellten Methoden nicht möglich ist, da die Energiewerte der paarweisen Wechselwirkung E_{ij} von geladenen Fragmenten um Größenordnungen höher sind als die der paarweisen Wechselwirkungsenergien der ungeladenen Fragmente. Die langreichweitigen Coulombwechselwirkungen überlagern die um Potenzen schwächeren Van-der-Waals-Wechselwirkungen. Letztere sind aber entscheidend für eine physikalisch richtige Beschreibung des chemischen Ensembles in MFD-Simulationen, insbesondere wenn ungeladene Fragmente an den Wechselwirkungen beteiligt sind. Damit DPD-Repulsionsparameter geladener Fragmente trotzdem innerhalb der molekularen Fragmentdynamik Anwendung finden, wurden im Rahmen dieser Arbeit folgende Annahmen entwickelt:

Der ungeladenen DPD-Repulsionsparameter $a_{i(uncharged),j(uncharged)}$ wird mittels

der MD-Boxmethode aus Kapitel 3.1.3 bestimmt.

Die Berechnung der ungleich geladenen DPD-Repulsionparameter $a_{i(charged),j(uncharged)}$ erfolgt unter Annahme der Bedingung, dass ein geladenes Fragment geringer repulsiv im Vergleich zu Wasser (H_2O) ist als das korrespondierende ungeladene Fragment:

$$
\begin{aligned}
&a_{i(charged),H_2O} = a_{i(uncharged),H_2O} - \Delta_i \\
&a_{\substack{j(uncharged),H_2O \\ i\neq j}} > a_{H_2O,H_2O} : \\
&\qquad a_{i(charged),j(uncharged)} = a_{i(uncharged),j(uncharged)} + \Delta_i \\
&a_{\substack{j(uncharged),H_2O \\ i\neq j}} \leqslant a_{H_2O,H_2O} : \\
&\qquad a_{i(charged),j(uncharged)} = a_{i(uncharged),j(uncharged)} - \Delta_i \\
&\text{mit } \Delta_k = \alpha \mid a_{i(uncharged),H_2O} - a_{H_2O,H_2O} \mid > 0;\ 0 \leq \alpha \leq 1
\end{aligned} \tag{3.4}
$$

Für gleich geladene DPD-Repulsionsparameter $a_{i(charged),j(charged)}$, wenn $q_i q_j > 0$, wobei q_i die Formalladung des Fragmentes i ist, gilt die Annahme, dass die Repulsion von gleich geladenen Fragmenten $q_i = q_j$ größer ist als jegliche andere Repulsion im betrachtetem Set.

$$a_{(charged),(charged)} = a_{max} + \Delta,\ \text{mit } \Delta = c_{correct}(a_{max} - a_{min}) > 0 \tag{3.5}$$

wobei a_{max} der maximale a_{ij} Parameter des Sets, a_{min} der minimale a_{ij} Parameter und $c_{correct} = 0.1 \mid q_i q_j \mid > 0$ ein beliebiger Ladungskorrekturfaktor ist. Für die Repulsion von ungleich geladenen Fragmenten $q_i \neq q_j$ gilt, dass sie kleiner ist als jede andere Repulsion im aktuellen Fragmentset:

$$a_{(charged),(charged)} = a_{min} - \Delta,\ \text{mit } \Delta = c_{correct}(a_{max} - a_{min}) > 0 \tag{3.6}$$

Abbildung 3.14 zeigt das mit dieser Methode vervollständigte Fragmentset inklusive aller geladenen Molkülfragmente. $c_{correct}$ wird bei der Erstellung mit einem Wert von 0.1 angewendet. Entsprechend der hier definierten Konvention werden positiv geladene Fragmente mit einem `P` und negativ geladene Fragmente mit

einem N als Suffix versehen.

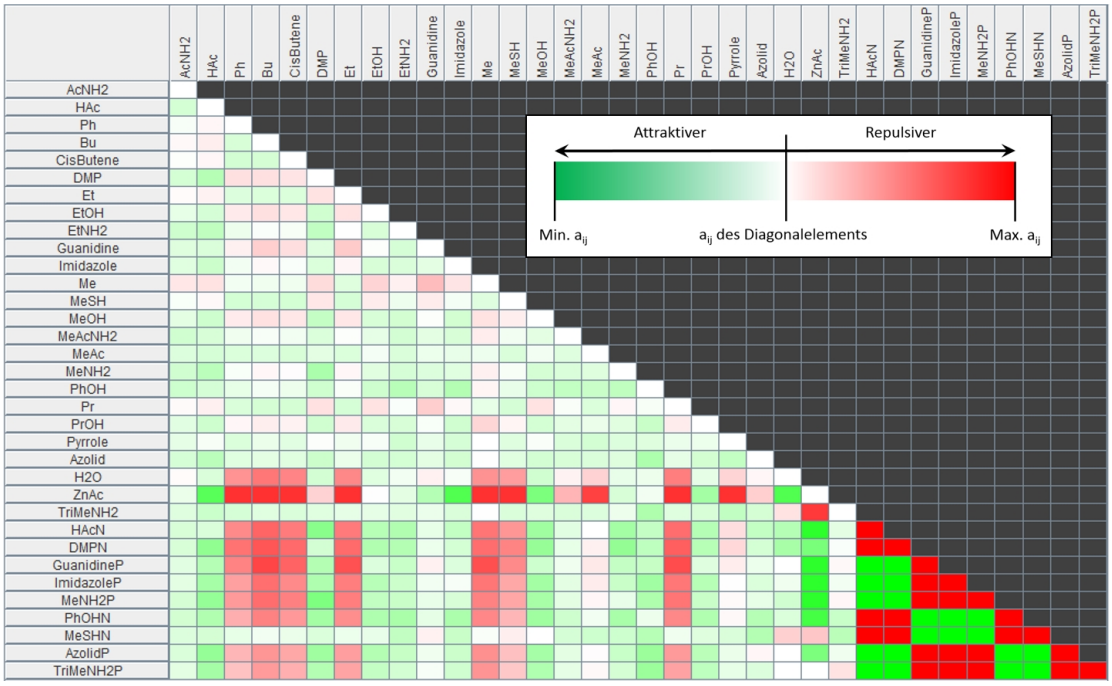


Abbildung 3.14: *Darstellung der DPD-Repulsionsparameter des MFD-Fragmentsets inklusive aller geladenen Fragmente.*

3.1.5 Einführung der elektrostatischen Kraft F^E in MFD-Simulationen

Elektrostatische Wechselwirkungen wurden für DPD von Groot [163] und Alejandre et al. [67] eingeführt. Die erste Methode verwendet eine Gittertechnik, um die Berechnung des elektrostatischen Feldes zu lösen. Die zweite Methode basiert auf der Ewald-Summation. Beide Methoden ersetzen die Punktladung eines Partikels durch eine verschmierte Ladungsverteilung, um die künstliche Anhäufung von Ionen zu verhindern [164].

Im Rahmen dieser Arbeit soll neben den ursprünglichen Kräften (konservative Kraft F^C, dissipative Kraft F^D und stochastische Kraft F^R) einer DPD-Simulation, die elektrostatische Kraft F^E zur Beschreibung von Coulomb-Ladungen innerhalb von MFD-Simulationen etabliert werden. Dabei wird die elek-

trostatische Kraft definiert als:

$$\underline{F}_{ij}^{E} = \frac{q_i q_j}{r_{ij}^2} \underline{r}_{ij}^0, \text{ mit } \underline{r}_{ij} = \underline{r}_i - \underline{r}_j;\ r_{ij} = \mid \underline{r}_{ij} \mid;\ \underline{r}_{ij}^0 = \frac{\underline{r}_{ij}}{r_{ij}} \tag{3.7}$$

Zur Erhöhung der Flexibilität der elektrostatischen Kraft soll eine effektive Ladung $q^* = c_{effective} \cdot q_i$, wobei $c_{effective}$ einen beliebigen Korrekturfaktor entspricht und die Effektivität der langreichweitigen Wirkung $r_{ij}^2 \rightarrow r_{ij}^k$, wobei k einen beliebigen Exponenten darstellt, eingesetzt werden. Eine gedämpfte Version der elektrostatischen Kraft für die in der molekularen Fragmentdynamik vorliegenden *soft-cores* ist zu definieren als:

$$\underline{F}_{ij}^{E} = \begin{cases} \dfrac{q_i^* q_j^*}{r_{ij}^k} \underline{r}_{ij}^0 & r_{ij} > r_{damping} \\ c_{damping} \dfrac{q_i^* q_j^*}{r_{ij}^k} \underline{r}_{ij}^0 & r_{ij} \leq r_{damping} \end{cases} \tag{3.8}$$

wobei $r_{damping}$ ein beliebiger Dämpfungsabstand und $c_{damping}$, mit $0 \leq c_{damping} \leq 1$ der Dämpfungsfaktor ist. Eine sinnvolle Wahl für $r_{damping}$ in MFD-Simulationen ist gleich dem Grenzwert für die konservative Kraft F^C. Zur Vermeidung von Singularitäten in der Anwendung der elektrostatischen Kraft kann zusätzlich eine maximale elektrostatische Kraft F_{max}^E ergänzt werden, welche gewählt wird, wenn $\underline{F}_{ij}^E \geq F_{max}^E$ gilt. Desweiteren wird ein Grenzwert r_{cutoff} zur Begrenzung der Reichweite eingeführt. Die elektrostatische Kraft ist vollständig definiert als:

$$\underline{F}_{ij}^{E} = \begin{cases} c_{damping} \dfrac{q_i^* q_j^*}{r_{ij}^k} \underline{r}_{ij}^0 & \mid \underline{F}_{ij}^E \mid < F_{max}^E; r_{ij} \leq r_{damping}; r_{ij} < r_{cutoff} \\ \dfrac{q_i^* q_j^*}{r_{ij}^k} \underline{r}_{ij}^0 & \mid \underline{F}_{ij}^E \mid < F_{max}^E; r_{ij} > r_{damping}; r_{ij} < r_{cutoff} \\ F_{max}^E & \mid \underline{F}_{ij}^E \mid \geq F_{max}^E; r_{ij} < r_{cutoff} \\ 0 & r_{ij} \geq r_{cutoff} \end{cases}, \tag{3.9}$$

3.2 Softwaretechnische Entwicklungen für die Simulation von biologischen Systemen

Für die Simulation von biologischen Systemen ist die Entwicklung von Software notwendig, die folgende Aufgaben übernimmt: (1) Bereitstellung und Aufbereitung von Peptid- und Proteindaten zur Verwendung in MFD-FormulaOne, (2) Visualisierung der molekularen und fragmentbasiereten Systeme sowie (3) Berechnung der DPD-Repulsionsparameter und die dazugehörige Verwaltung des Fragmentsets.

Aufgabe (1) wird von der in MFD-FormulaOne vollständig integrierten Anwendung `MfdPeptide` übernommen. Sie dient als Schnittstelle zwischen der Welt der Biopolymere wie Peptiden und Proteinen und der Simulationssoftware MFD-FormulaOne. Das Programm gliedert sich in zwei Teile:

1. Verwaltung der Aminosäuredaten, Konvertierung von Aminosäuresequenzen in einen *f*SMILES sowie dessen automatisierte Ladung in Abhängigkeit von einem gegebenen pH-Wert (Kapitel 3.2.1).

2. Import der Proteinsequenz und geometrischen Informationen aus *Protein-Data-Bank*-(PDB)-Archivdateien[165], Anwendung von geometrischen Operationen (Translation, Rotation und Skalierung) auf die vorliegenden Daten sowie Bereitstellung aller Funktionen zur Nutzung der Stabilisierungspotentiale des Proteinrückgrats (Kapitel 3.2.2).

Die Visualisierung der Proteindaten sowie der fragmentbasierten Simulationsboxen von Punkt (2) übernimmt in MFD-FormulaOne die Anwendung `MfdJmolViewer` (Kapitel 3.2.3). Sie dient als Adapter zwischen MFD-FormulaOne und dem *Open-Source*-Projekt *Jmol*[47].

Aufgabe (3) übernehmen die Anwendungen `MfdCompassFX` (Kapitel 3.2.4) und `MfdFragmentManager` (Kapitel 3.2.5). `MfdCompassFX` stellt dem Nutzer eine grafische Benutzeroberfläche mit allen notwendigen Funktionen zur Berechnung von DPD-Repulsionsparametern für MFD-Simulationen zur Verfügung. Das Programm `MfdFragmentManager` dient der Verwaltung aller Eigenschaften

der Molekülfragmente, der dazugehörigen DPD-Repulsionsparameter sowie der notwendigen Daten für die Verwendung von Aminosäuren in der molekularen Fragmentdynamik.

Alle Anwendungen wurden in Java in der Version 1.7 (`MfdCompassFX` in der Version 1.8) entwickelt. Die Anwendung `MfdFragmentManager` verwendet die Programmierschnittstelle (engl. *Application Programming Interface*, kurz API) *Java Swing* zur Darstellung der grafischen Benutzeroberfläche, die Anwendung `Mfd-CompassFX` nutzt *JavaFX*. Zusammen umfassen die Anwendungen 14690 Codezeilen bzw. 19174 Codezeilen inklusive Unittests.

3.2.1 Behandlung von Aminosäuresequenzen

Die nächsthöhere Hierarchieebene über den Molkülfragmenten zur Beschreibung von Peptiden und Proteinen ist die Aminosäure. Eine Aminosäure wird in der molekularen Fragmentdynamik anhand folgender Parameter vollständig beschrieben: Name, Ein- und Dreibuchstabencode, zugehöriger *f*SMILES sowie Kodierung ladungsspezifischer Eigenschaften. Abbildung 3.15 zeigt einen Auszug der Aminosäuresektion aus den Daten des hier entwickelten Fragmentsets. Im Anhang findet sich eine komplette Liste der aufgezählten Parameter für alle proteinogenen Aminosäuren (Tabellen 11.2 und 11.3)

```
[Amino acids]
# Amino acids for MFD FormulaOne
# Columns: One-letter code / Three-letter code / Name / fSmiles / Charges
A Ala Alanine       MeAcNH2BB(Me)            TC&2.3?HAcBB:HAcNBB-;TN&9.9?MeNH2PBB+:MeNH2BB;
R Arg Arginine      MeAcNH2BB(Pr-Guanidine)  TC&2.81?HAcBB:HAcNBB-;TN&9.09?MeNH2PBB+:MeNH2BB;
                                             SCF&12.48?GuanidineP+:Guanidine;
                                             SCB&12.48?GuanidineP+:Guanidine;
N Asn Asparagine    MeAcNH2BB(AcNH2)         TC&2.02?HAcBB:HAcNBB-;TN&8.8?MeNH2PBB+:MeNH2BB;
D Asp AsparticAcid  MeAcNH2BB(HAc)           TC&1.88?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2BB;
                                             SCF&3.68?HAc:HAcN-;SCB&3.85?HAc:HAcN-;
C Cys Cysteine      MeAcNH2BB(MeSH)          TC&1.71?HAcBB:HAcNBB-;TN&10.78?MeNH2PBB+:MeNH2BB;
                                             SCF&8.33?MeSH:MeSHN-;SCB&8.95?MeSH:MeSHN-;
                                             SS&MeSH?MeSHSS;
E Glu GlutamicAcid  MeAcNH2BB(Me-HAc)        TC&2.19?HAcBB:HAcNBB-;TN&9.67?MeNH2PBB+:MeNH2BB;
                                             SCF&4.25?HAc:HAcN-;SCB&4.35?HAc:HAcN-;
Q Gln Glutamine     MeAcNH2BB(Me-AcNH2)      TC&2.17?HAcBB:HAcNBB-;TN&9.13?MeNH2PBB+:MeNH2BB;
G Gly Glycine       MeAcNH2BB                TC&2.21?HAcBB:HAcNBB-;TN&9.15?MeNH2PBB+:MeNH2BB;
H His Histidine     MeAcNH2BB(Me-Imidazole)  TC&1.78?HAcBB:HAcNBB-;TN&8.97?MeNH2PBB+:MeNH2BB;
                                             SCF&6.0?ImidazoleP+:Imidazole;
                                             SCB&6.9?ImidazoleP+:Imidazole;
I Ile Isoleucine    MeAcNH2BB(Et-Et)         TC&2.32?HAcBB:HAcNBB-;TN&9.76?MeNH2PBB+:MeNH2BB;
L Leu Leucine       MeAcNH2BB(Me-Pr)         TC&2.4?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2BB;
K Lys Lysine        MeAcNH2BB(Pr-MeNH2)      TC&2.2?HAcBB:HAcNBB-;TN&8.9?MeNH2PBB+:MeNH2BB;
                                             SCF&10.53?MeNH2P+:MeNH2;SCB&9.4?MeNH2P+:MeNH2;
M Met Methionine    MeAcNH2BB(Et-MeSH)       TC&2.28?HAcBB:HAcNBB-;TN&9.21?MeNH2PBB+:MeNH2BB;
F Phe Phenylalanine MeAcNH2BB(Me-Ph)         TC&2.58?HAcBB:HAcNBB-;TN&9.24?MeNH2PBB+:MeNH2BB;
P Pro Proline       AzolidBB                 TC&1.99?HAcBB:HAcNBB-;TN&10.6?AzolidPBB+:AzolidBB;
S Ser Serine        MeAcNH2BB(MeOH)          TC&2.21?HAcBB:HAcNBB-;TN&9.15?MeNH2PBB+:MeNH2BB;
T Thr Threonine     MeAcNH2BB(PrOH)          TC&2.1?HAcBB:HAcNBB-;TN&9.12?MeNH2PBB+:MeNH2BB;
W Trp Tryptophan    MeAcNH2BB(Me-Pyrrole-Ph) TC&2.15?HAcBB:HAcNBB-;TN&9.12?MeNH2PBB+:MeNH2BB;
Y Tyr Tyrosine      MeAcNH2BB(Me-PhOH)       TC&2.2?HAcBB:HAcNBB-;TN&9.11?MeNH2PBB+:MeNH2BB;
                                             SCF&10.07?PhOH:PhOHN-;SCB&9.9?PhOH:PhOHN-;
V Val Valine        MeAcNH2BB(Pr)            TC&2.3?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2BB;
[/Amino acids]
```

Abbildung 3.15: *Ausschnitt der Definitionsdatei für das Fragmentset aller 20 proteinogenen Aminosäuren für die Verwendung in MFD-FormulaOne.*

Die Syntax der Ladungskodierung basiert auf der Syntax des Bedingungsoperators: „<Bedingung> ? <Ausdruck> : <Ausdruck>“. Als Beispiel soll die Ladungskodierung für die Ladung der C-terminalen Carboxylgruppe von Alanin dienen: TC&2.3?HAcBB:HAcNBB-. Die Bedingung setzt sich immer aus zwei Parametern zusammen, welche durch ein '&' Zeichen voneinander getrennt sind. Der erste Teil bestimmt das zu manipulierende Fragment. Folgende Werte sind möglich:

- TC: Carboxylgruppe der Peptidbindung.
- TN: Aminogruppe der Peptidbindung.
- SCF: Seitenkette, wenn die Aminosäure ungebunden vorliegt.
- SCB: Seitenkette, wenn die Aminosäure gebunden vorliegt.

Der zweite Parameter, der pKs-Wert der korrespondierenden funktionellen Gruppe der Aminosäure, beschreibt das zu verwendende Fragment (erster Ausdruck oder zweiter Ausdruck) bezüglich des gegebenen pH-Wertes. Ist der gegebene pH-Wert kleiner als der vorgegebene pKs-Wert, wird das Fragment des ersten Ausdruckes eingesetzt. Dieser ist als die protonierte Form definiert. Ist der gegebene pH-Wert größer als der angegebene Wert, so wird die deprotonierte Form des zweiten Ausdruckes verwendet. In unserem Beispiel wird bei einem pH-Wert von 7.0 folgender *f*SMILES für Alanin (ungeladen: `MeNH2BB(Me)-HAcBB`) generiert: `MeNH2PBB(Me)-HAcNBB`.

Dieser Ansatz weist zwei notwendige Vereinfachungen auf: (1) Die mit dem hier abgebildeten System bestimmte Verteilung zwischen geladenen zu ungeladenen Fragmenten ist sprunghaft, d.h. es sind entweder alle Fragmente geladen bzw. ungeladen. Die Verteilung entspricht nicht dem sigmoiden Verlauf einer Gleichgewichtsreaktion. Dies führt bei geringen Unterschieden vom angegeben pH-Wert zu den einzelnen pKs-Werten zu falschen Ladungsverteilungen. (2) Die pKs-Werte der spezifischen funktionellen Gruppen der Aminosäuren sind nicht gleich innerhalb eines Proteins, sondern abhängig von der chemischen Umgebung. Die hier verwendeten pKs-Werte entsprechen den durchschnittlichen Werten und weichen so von den wahren Werten ab. Jedoch werden alle Simulationen bei einem physiologischen pH-Wert von 7.4 durchgeführt. Dieser ist ausreichend weit vom nächsten pKs-Wert mit 6.9 (Histidinseitenkette) entfernt, sodass es nur zu einem geringen Fehler bezüglich der Ladungsverteilung kommt.

Als weiteres Beispiel soll die Konvertierung des Pentapeptides `RRHIS` dienen. Das Peptid besitzt folgenden *f*SMILES:

```
MeNH2BB(Pr-Guanidine)
-MeAcNH2BB(Pr-Guanidine)
-MeAcNH2BB(Me-Imidazole)
-MeAcNH2BB(Et-Et)
-MeAcNH2BB(MeOH)-HAcBB
```

Für eine adäquate Beschreibung des Ladungsstatus der Seitenkette und von terminalen funktionalen Gruppen in Abhängigkeit des pH-Wertes wird die Definition

der Aminosäuresequenz im Einbuchstabencode erweitert. Beispielsweise besitzt das Musterpeptid `RRHIS` bei einem pH-Wert von 7.4 einen positiv geladenen N-Terminus, zwei positiv geladene Histidinseitenketten und einen negativ geladenen C-Terminus. Dies kann durch die Einführung von zusätzlichen Symbolen R{N+S+}R{S+}HIS{C-} gekennzeichnet werden. {N+} beschreibt hierbei den positiv geladenen N-Terminus und {C-} den negativ geladenen C-Terminus. {S-} oder {S+} kennzeichnen jeweils eine negativ bzw. positiv geladene Seitenkette. Hier im Beispiel existieren zwei positiv geladene Argininseitenketten, bestehend aus je einem Propan (Pr) und einem geladenen Guanidinium (GuanidinP) Fragment:

```
MeNH2PBB(Pr-GuanidineP)
-MeAcNH2BB(Pr-GuanidineP)
-MeAcNH2BB(Me-Imidazole)
-MeAcNH2BB(Et-Et)
-MeAcNH2BB(MeOH)-HAcNBB
```

Abbildung 3.16 zeigt die Implementierung in die grafische Benutzeroberfläche (engl. *Graphical User Interface*, kurz GUI) von MFD-FormulaOne. Das Textfeld *Peptide* (1) ermöglicht die Eingabe der Aminosäuresequenzen im Einbuchstabencode. Darunter befindet sich die Statuszeile zur Darstellung der Validität der eingegebenen Aminosäuresequenz (2). Das unterste Textfeld zeigt die eingegebene Aminosäuresequenz im Dreibuchstabencode (3). Auch die Definition des Dreibuchstabencodes ist äquivalent zum Einbuchstabencode um Ladungseigenschaften erweitert. Die Schaltfläche *pH* (4) öffnet einen Dialog zur Eingabe des pH-Wertes. Die Bestätigung des Dialogs führt zur automatischen Ladung der vorhandenen Aminosäuresequenz. Die Auswahlliste (5) gibt dem Nutzer neben einer Übersicht aller verfügbaren Aminosäuren eine Liste der verfügbaren Sonderelemente wie Disulfidbrückenbindungen, Ladungsinformationen sowie die Definition ringförmiger Proteinrückgrate (6).

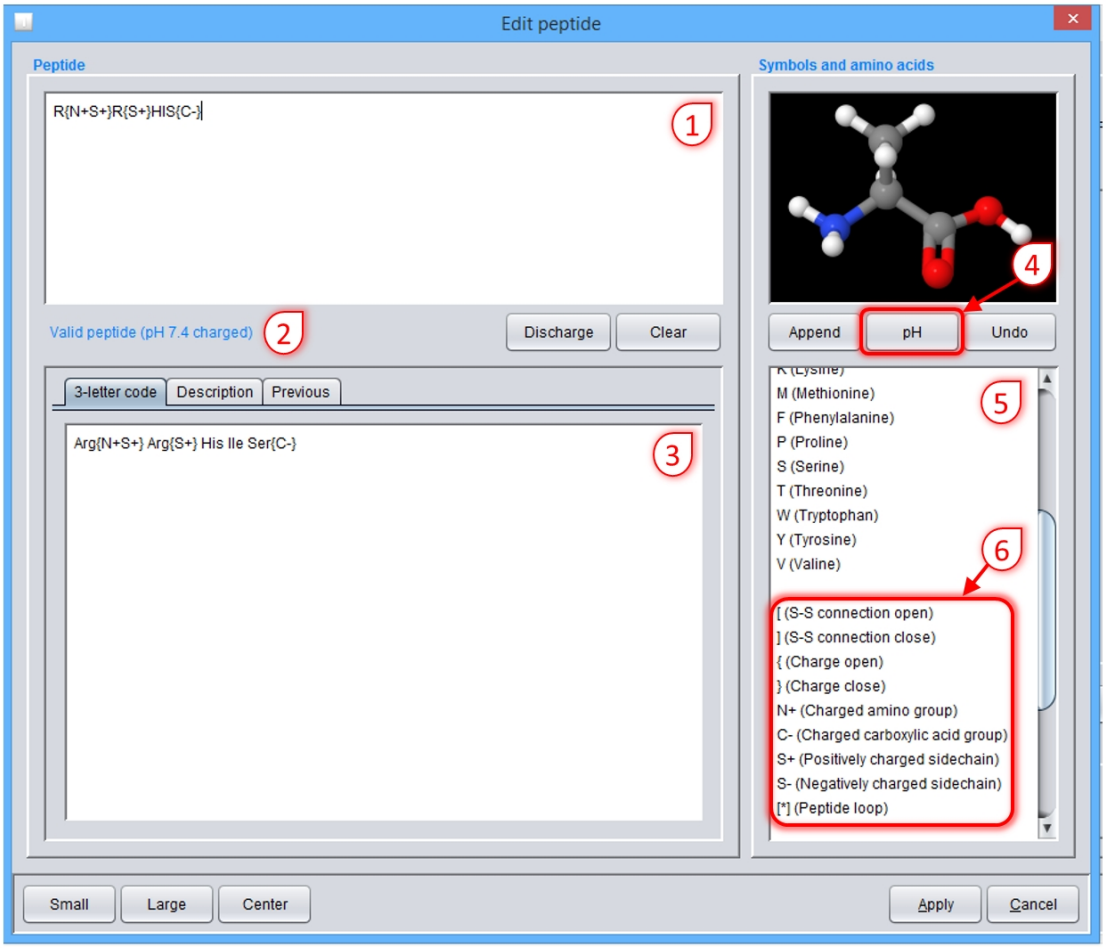


Abbildung 3.16: *Dialog zur Eingabe und Bearbeitung von Aminosäuresequenzen in MFD-FormulaOne.*

Bei Bestätigung der Eingabe werden die Daten automatisch in einen *f*SMILES überführt. Das übergeordnete Fenster *Edit molecular structure* übernimmt den generierten *f*SMILES und zeigt einen ähnlichen Aufbau wie der *Edit peptide* Dialog. Das obere Textfeld (Abbildung 3.17, 1) zeigt den *f*SMILES. Unten ist eine grafische Darstellung der Fragmenttopologie dargestellt (2). Ermöglicht wird diese mit dem *Open-Source*-Projekt *Chemistry Development Kit* (CDK)[166,167].

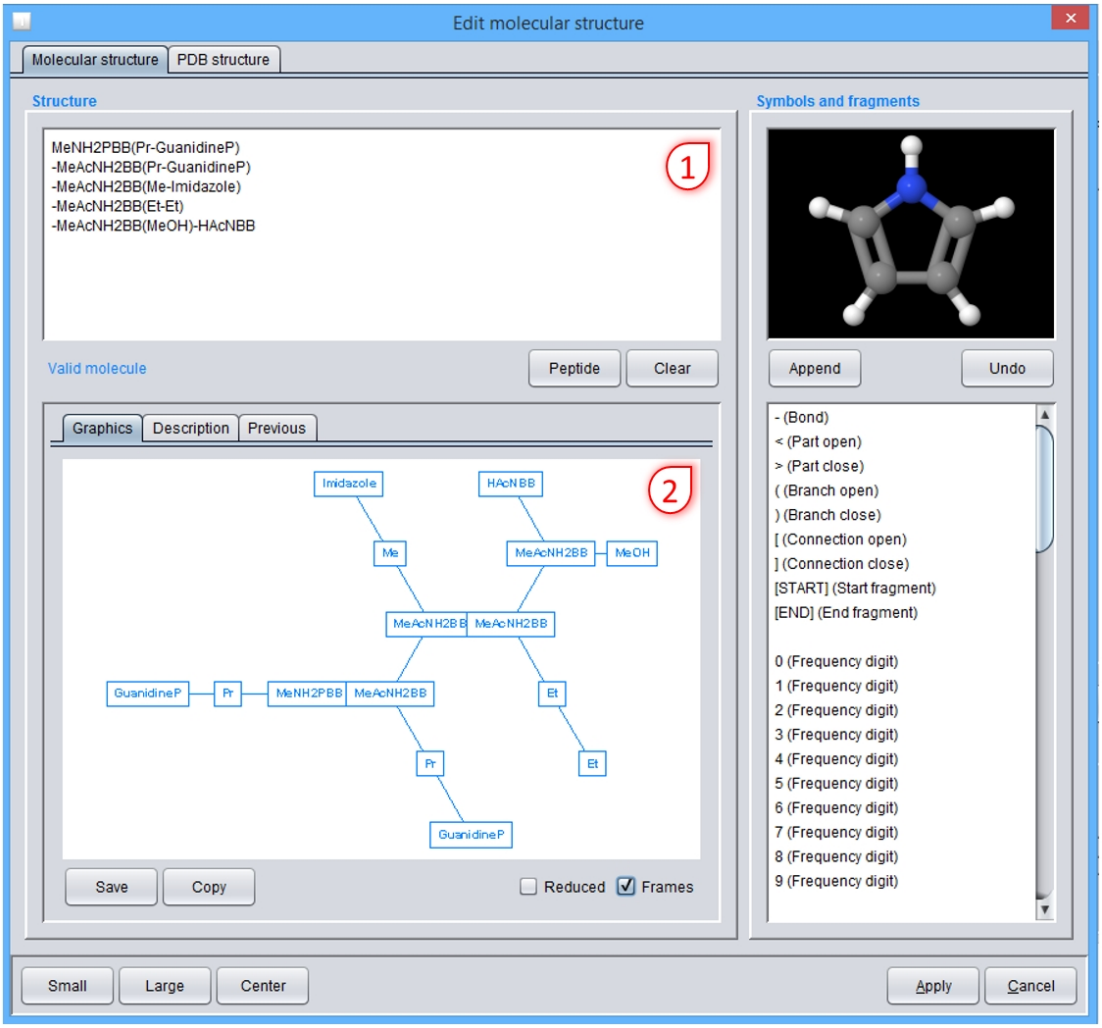


Abbildung 3.17: *Darstellung der konvertierten Aminosäuresequenz* ***RRHIS*** *im Moleküleditierungsdialog von MFD-FormulaOne.*

3.2.2 Unterstützung für Protein Data Bank (PDB)-Archivdateien

Ein durchschnittliches Protein besitzt eine Sequenzlänge von 316 Aminosäuren[45]. Abbildung 3.18 zeigt die Verteilung der Sequenzlängen in den Datenbanken UniProtKB/TrEMBL mit insgesamt $8.6 \cdot 10^7$ Einträgen.

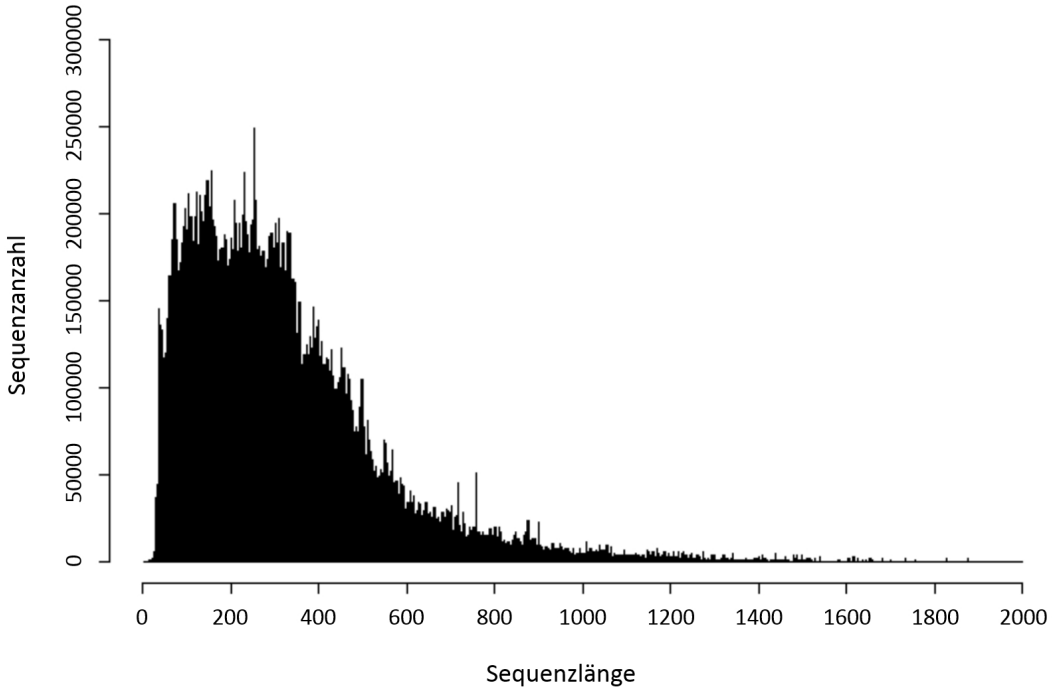


Abbildung 3.18: *Verteilung der Länge der Aminosäuresequenz in der Datenbank UniProtKB/TrEMBL mit* $8.67 \cdot 10^7$ *Einträgen*[45].

Dies zeigt die Notwendigkeit der Entwicklung eines Werkzeuges zur automatisierten Generierung von für die Anwendung MFD-FormulaOne interpretierbaren Proteindaten. Neben der Beschreibung der Proteine in Form eines *f*SMILES soll zusätzlich die dreidimensionale Konformation der Proteine mit der Topologie verknüpft werden, um ein dreidimensionales Abbild des Proteins in MFD-Simulationen verwenden zu können. Das Programm `MfdPeptide` ermöglicht die Extraktion der Konformationsinformationen sowie der Sequenzdaten aus einer PDB-Archivdatei[165,168] und bereitet diese für die Verwendung in MFD-FormulaOne auf.

Die Oberfläche *PDB structure* (siehe Abbildung 3.19) des *View protein*-Dialogs zeigt die Einbettung von `MfdPeptide` in MFD-FormulaOne für die Unterstützung von PDB-Archivdateien. Das Laden von gespeicherten PDB-Archivdateien übernimmt die Schaltfläche *Load* (1). Zum Zweck der Visualisierung der geladenen Proteindaten wird das Programm `MfdJmolViewer` auf Basis von Jmol[47] (siehe Kapitel 3.2.3) eingebunden (2). Neben der Oberfläche *3D structure* sind drei weitere Ansichten der Proteindaten wählbar: *Sequence(s)*, *Molecular structure* und

PDB file (3). Unter *Sequence(s)* verbirgt sich die Ansicht der um Ladungen erweiterten Aminosäuresequenz und *Molecular structure* stellt den aktuellen *f*SMILES des Proteins dar. *PDB file* zeigt die zugrunde liegende PDB-Archivdatei in Textform an.

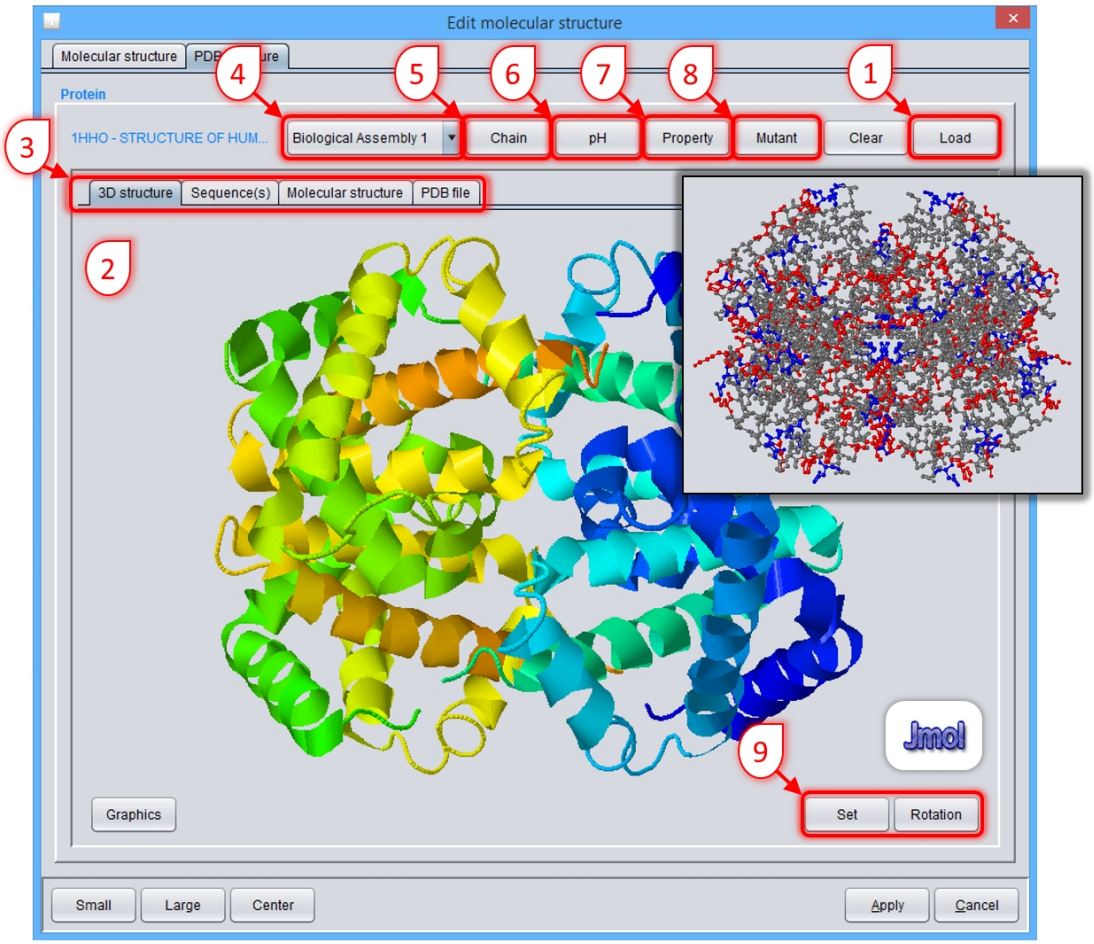


Abbildung 3.19: *Die Oberfläche PDB structure des Edit molecular structure Dialogs. Das kleine Bild zeigt die geladenen Aminosäuren des Proteins Hämoglobin (PDB-ID: 1HHO[169]) bei einem pH-Wert von 7.4 (Grau: Neutral, rot: Positiv, blau: Negativ).*

Die Auswahlbox (4) ermöglicht die Auswahl verschiedener biologischer Anordnungen der asymmetrischen Proteinuntereinheiten, definiert unter REMARK 350 in der PDB-Datei. Die Schaltfläche *Chain* (5) öffnet einen Dialog der zwei Einstellungen ermöglicht: Die erste Einstellung *Chain selection* erlaubt es dem Nutzer, gezielt Proteinketten ein- bzw. auszuschalten, sodass einzelne Ketten je nach gesetztem Status von der Simulation ausgeschlossen werden können. Die zweite

Option *Chain segment assignment* weist bei Aktivierung dem Proteinrückgratsegmenten (Erklärung siehe unten) jeder Proteinkette eine eigene Nummer zu. Der Standardwert ist 0. Mittels der Schaltfläche (6) lässt sich der pH-Wert für die entsprechende Ladung des Proteins bestimmen. Die Auswirkung für den physiologischen pH-Wert von 7.4 ist im kleinen Bildausschnitt in Abbildung 3.19 für das Protein Hämoglobin (PDB-ID: 1HHO[169]) zu sehen (Grau: Neutral, rot: Positiv, blau: Negativ). Die Schaltfläche *Property* (7) öffnet einen Dialog, der drei Einstellungen ermöglicht: *Backbone probes*, *Backbone fragment status* und *Sequence loop*. Unter dem Punkt *Backbone probes* ist es möglich, den Austausch von α-C-Fragmenten durch Sondenfragmente zu definieren. Das Einfügen von Sonden ermöglicht die Betrachtung von räumlichen Änderungen innerhalb der Proteingeometrie während der Simulation. Die Einstellung *Backbone fragment status* ist ausführlich im weiteren Verlauf des Kapitels erklärt. Die Eigenschaft *Sequence loop* produziert bei Aktivierung einen geschlossenen Ring des Proteinrückgrates, d.h. die beiden Termini (N-Terminus und C-Terminus) werden miteinander kovalent verknüpft (siehe Kapitel 7.4). Unter dem Punkt *Mutant* (8) können Mutationen durch den Austausch einzelner Aminosäuren durch andere erzeugt werden. Mit den Schaltflächen *Set* und *Rotation* lässt sich die räumliche Ausrichtung aus der Ansicht des Proteins (2) für die Simulationen übernehmen.

Als Startgeometrie eine zufällige Startkonfiguration zu verwenden, kann für spezifische Simulationen von Nachteil sein und vermeidbar lange Simulationszeiten erfordern, bis das thermodynamische Gleichgewicht eingestellt ist. Damit eine Startgeometrie mit höherer Ordnung erstellt werden kann, liefert MFD-FormulaOne sogenannte Kompartimente (engl. *Compartments*) in Form von Kugeln (Sphären), Schichten und Ziegeln, die frei in der Simulationsbox platziert werden können.

Proteine mit Atompositionsdaten können in Form ihrer nativen Konformation in einem Kugelkompartiment untergebracht werden. Zusätzlich können mehrere Proteinstrukturen zufällig in einem Schichtkompartiment platziert werden. Die räumliche Anordnung der Fragmente ist nur für α-C-Fragmente exakt spezifiziert, d.h. diese werden korrekt auf die korrespondierenden Koordinaten der α-C-Atome

der Kristallstruktur des Proteins gesetzt. Die restlichen Fragmente der jeweiligen Seitenkette übernehmen dieselbe Raumkoordinate. Die Relaxation des Systems erfolgt bei Simulationsbeginn. Die Ausrichtung der Seitenkettenfragmente wird nur über die Simulationsdynamik bestimmt. Für die Platzierung der Proteinfragmente innerhalb der Sphären- und Schichtkompartimente bietet *MfdPeptide* verschiedene Methoden der Koordinatentransformation (Translation, Rotation und Skalierung). Abbildung 3.20 zeigt eine Simulationsbox (rechts) mit der Anordnung eines G-Protein-gekoppelten Rezeptors (PDB-ID: 2RH1 [170], grün) unter einer Phosphatidylethanolaminmembran. Das G-Protein wird mittels eines Sphärenkompartiments direkt unter dem Schichtkompartiment der Membran platziert.

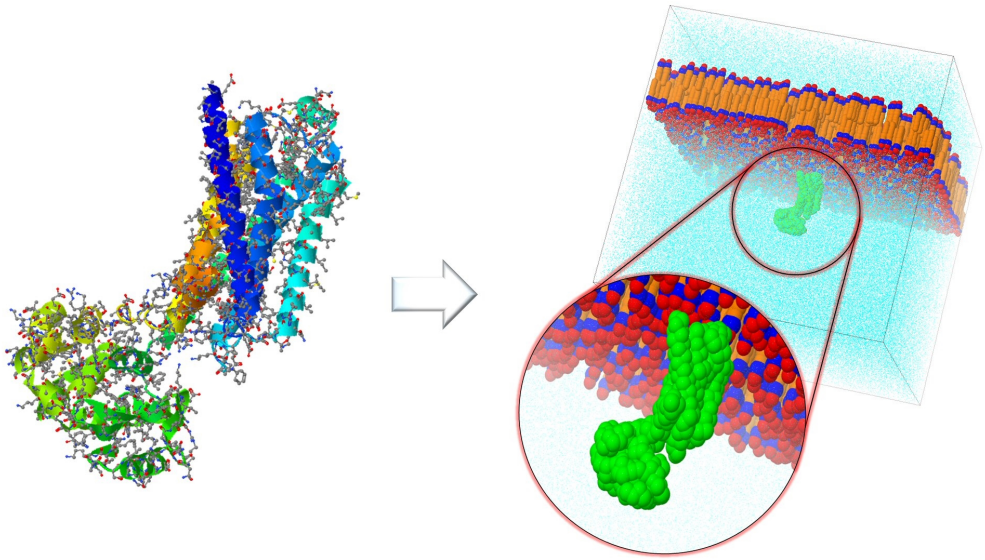


Abbildung 3.20: *Startgeometrie (rechts) eines G-Protein-gekoppelten Rezeptors (PDB-ID: 2RH1 [170]; Cartoon-Darstellung mit eingeblendeten Seitenketten links, grün rechts) unter einer Phosphatidylethanolaminmembran.*

Mittels der Fragmenttopologie in Form eines *f*SMILES, der räumlichen Beschreibung der Proteingeometrie und den Gestaltungsmöglichkeiten der Startgeometrie der Simulationsbox anhand von Kompartimenten können flexibel verschiedenste Anfangskonfigurationen entworfen werden. Wie in Kapitel 7.1.1 gezeigt wird, kann die Stabilisierung der Proteinstruktur notwendig sein. Für die Anwendung der in Kapitel 7.1 beschriebenen Konformationsstabilisierung müssen die entsprechenden

α-C-Fragmente für die Verwendung mit den harmonischen Potentialen gekennzeichnet werden. Abbildung 3.21 zeigt einen Auszug der bereitgestellten Daten für die Auswertung der anzuwendenden harmonischen Potentiale durch die MFD-Engine. Jedem α-C-Fragment wird ein eindeutiger Potentialindex zugewiesen. Mit einer 0 gekennzeichnete Fragmente werden ignoriert. Des Weiteren ist aus der Abbildung ersichtlich, dass nicht-α-C-Fragmente dieselben Koordinaten wie die jeweiligen α-C-Fragmente zugeordnet bekommen. Die letzte Spalte stellt die Bindungen der einzelnen Fragmente dar. Ein negatives Vorzeichen kennzeichnet eine Verknüpfung zu einem voranstehenden Fragment, ein positives Vorzeichen ein nachfolgendes Fragment. Der Zahlenwert beschreibt die Differenz des Indices zum Bindungspartner.

```
#-----------------------------------------------------------------------
Number_Of_Fragments 336
Fragments_Per_Molecule 336
#-----------------------------------------------------------------------
# Compartment Sphere with Protein
#-----------------------------------------------------------------------
# Compartment Sphere Protein Information:
# Protein Name              = Calmodulin
# Fragments per Protein     = 336
# Quantity in Volume        = 1
# Quantity on Surface       = 0
#-----------------------------------------------------------------------
# Index Fragment    Potential-Index x       y       z       Bond-Offsets
#-----------------------------------------------------------------------
  1     MeAcNH2PD1  1               11.337  13.951  13.161   1  3
  2     Et          0               11.337  13.951  13.161  -1  1
  3     Et          0               11.337  13.951  13.161  -1
  4     MeAcNH2BB   2               11.075  13.873  13.384  -3  1   2
  5     Me          0               11.075  13.873  13.384  -1
  6     MeAcNH2BB   3               11.197  13.543  13.380  -2  1   3
  7     Me          0               11.197  13.543  13.380  -1  1
  8     HAcN        0               11.197  13.543  13.380  -1
  9     MeAcNH2BB   4               11.509  13.677  13.467  -3  1   3
 10     Me          0               11.509  13.677  13.467  -1  1
 11     Ph          0               11.509  13.677  13.467  -1
 12     MeAcNH2BB   5               11.358  13.883  13.715  -3  1   3
 13     Pr          0               11.358  13.883  13.715  -1  1
```

Molekülinformationen

Fragmentinformationen

...

Abbildung 3.21: *Beschreibung des topologischen und räumlichen Aufbaus des Proteins sowie die eindeutige Stabilisierungspotentialindizierung der α-C-Fragmente.*

An dieser Stelle sind die α-C-Fragmente für die Potentiale eindeutig indexiert und können verwendet werden. Jetzt müssen die Potentiale noch definiert und mit einer Kraftkonstante sowie mit der Distanz zwischen dem Fragmentpaar versehen

werden. Abbildung 3.22 zeigt einen Ausschnitt der ersten Potentiale mit einer Distanz von einem α-C-Fragment. Zusätzlich sind die mittlere Distanz des Fragmentpaares und die eingestellte Kraftkonstante angegeben. Die aus der Kristallstruktur abgeleiteten Distanzen mit einer Potentiallänge von einem α-C-Fragment entsprechen immer genau der mittleren Distanz zwischen zwei α-C-Atomen mit 3.8 ± 0.1 Å[171].

```
#-----------------------------------------------
# Index A   Index B Distance    Force-Constant
#-----------------------------------------------
  1          2        3.863       4.000
  2          3        3.777       4.000
  3          4        3.744       4.000
  4          5        3.724       4.000
  5          6        3.800       4.000
  6          7        3.786       4.000
  7          8        3.858       4.000
  8          9        3.785       4.000
  9         10        3.837       4.000
 10         11        3.856       4.000
 11         12        3.769       4.000
 12         13        3.804       4.000
 13         14        3.781       4.000
                        ...
```

Abbildung 3.22: *Tabelle mit Potentialdefinitionen.*

Abbildung 3.23 zeigt den Dialog in MFD-FormulaOne, in dem die zu verwendenden Potentiale in Abhängigkeit ihrer Distanz und die zugehörige Kraftkonstante einzustellen sind.

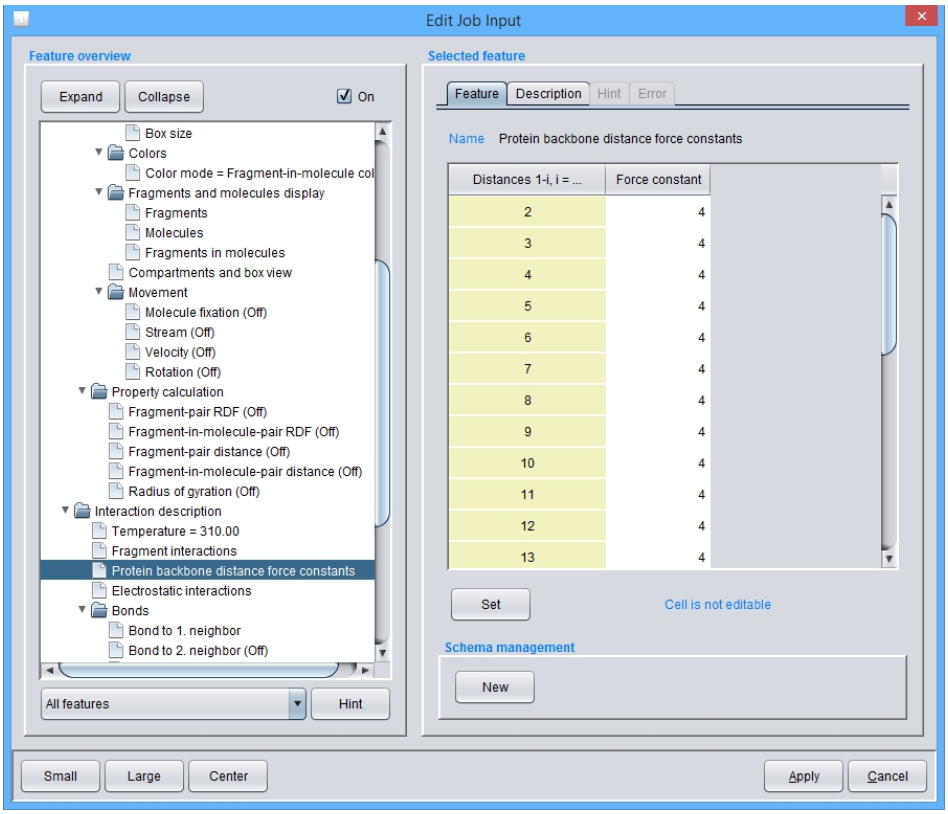


Abbildung 3.23: *Dialog zum Setzen der Stabilisierungspotentiale sowie der zu verwendenden Kraftkonstanten.*

Für eine bessere Kontrolle über die Stabilisierungspotentiale wurden zwei weitere Einstellungsmöglichkeiten implementiert: Der Proteinrückgratfragmentstatus und die Proteinrückgratfragmentsegmente. Diese Einstellungen sind über die Schaltfläche *Property* (siehe Abbildung 3.19, 7) unter dem Menüpunkt *Backbone fragment status* erreichbar. Mit dem Proteinrückgratfragmentstatus ist es möglich, gezielt Fragmente auszuschließen, sodass sie während des Erstellungsprozesses der Stabilisierungspotentiale ignoriert werden.

Mittels der Segmenteinstellung lassen sich die Fragmente des Proteinrückgrates in verschiedene nummerierte Segmente einteilen. Dies hat zur Folge, dass es keine Potentiale geben kann, welche über die Grenze zweier verschieden nummerierter Segmente wirken können. Diese Eigenschaft kann dazu verwendet werden, um beispielsweise zwei unterschiedliche Proteinketten getrennt voneinander zu behandeln (siehe Kapitel 7.2). Abbildung 3.24 zeigt die zugehörige Benutzeroberfläche, in welcher der Benutzer die beschriebenen Einstellungen vornehmen kann.

Name Backbone fragment status

Index	Amino acid	Status	Segment
1	Cysteine	Off	1
2	Glycine	Off	1
3	GlutamicAcid	On	1
4	Threonine	On	1
5	Cysteine	On	1
6	Valine	On	1
7	Glycine	On	2
8	Glycine	On	2
9	Threonine	On	2
10	Cysteine	On	2
11	Asparagine	Off	2
12	Threonine	Off	2

Abbildung 3.24: *Ausschnitt der MFD-FormulaOne Benutzeroberfläche zum Einstellen des Proteinrückgratfragmentstatus und der Proteinrückgratfragmentsegmente der α-C-Fragmentpotentiale.*

3.2.3 Visualisierung von molekularen und fragmentbasierten Systemen

Die Anwendung `MfdJmolViewer` dient als Adapter zwischen MFD-FormulaOne und dem *Open-Source*-Projekt *Jmol*[47]. *Jmol* ist ein Programm zur interaktiven Visualisierung von molekularen Strukturen, insbesondere von Biomolekülen. Die Anwendung ist bzgl. der räumlichen Darstellung von molekularen Vielteilchensystemen optimiert und so im Rahmen dieser Arbeit entsprechend anwendbar.

Der `MfdJmolViewer` bietet für MFD-FormulaOne zwei Einsatzmöglichkeiten: Die Präsentation von chemischen Strukturen wie Proteinen aus PDB-Archivdateien sowie der Darstellung der fragmentbasierten Simulationsboxen. Abbildung 3.25 zeigt die Implementierung zur Darstellung von Proteinen in MFD-FormulaOne und die beiden möglichen Visualisierungsstile: Kugel-Stab- und Cartoon-Modell. Beide Stile sind auch überlagert darstellbar.

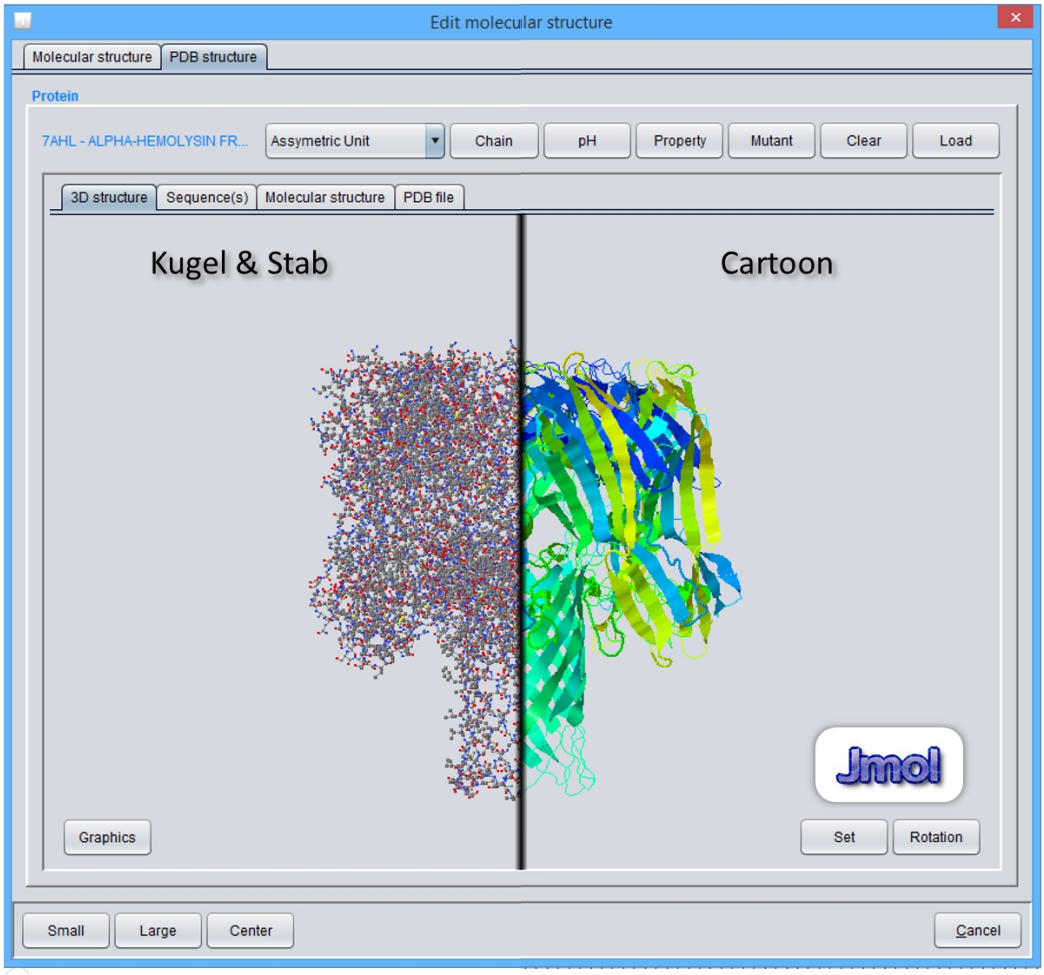


Abbildung 3.25: *Einbettung des* `MfdJmolViewers` *in die Anwendung MFD-FormulaOne zur Darstellung von Proteinen. Zu sehen ist das Protein α-Hämolysin (PDB-ID: 7AHL[172]) von Staphylococcus aureus (links: Kugel-Stab-Modell; rechts: Cartoon-Darstellung der Sekundärstruktur).*

Wenn das Kugel-Stab-Modell gewählt ist, sind vier weitere Färbungen des Biomoleküls verfügbar:

- Polarität (Polarity):

  Hervorhebung der hydrophilen (rot) und hydrophoben (grün) Aminosäuren.

- Ladung (Charge):

  Ladungstragende Aminosäuren werden farblich gekennzeichnet (rot: Positiv, blau: Negativ, grau: Neutral).

- Proteinrückgrat (Backbone):

Atome des Proteinrückgrates werden rot dargestellt, alle sonstigen Atome grau.

- α-Kohlenstoffatome:
  Nur α-Kohlenstoffatome der Aminosäuren sind sichtbar. Die Positionen der α-Kohlenstoffatome repräsentieren die spätere Lage der korrespondierenden Aminosäurefragmente in der Startgeometrie einer Simulation.

Abbildung 3.26 skizziert die möglichen Färbungen anhand des Proteins α-Hämolysin (PDB-ID: 7AHL[172]) von *Staphylococcus aureus*.

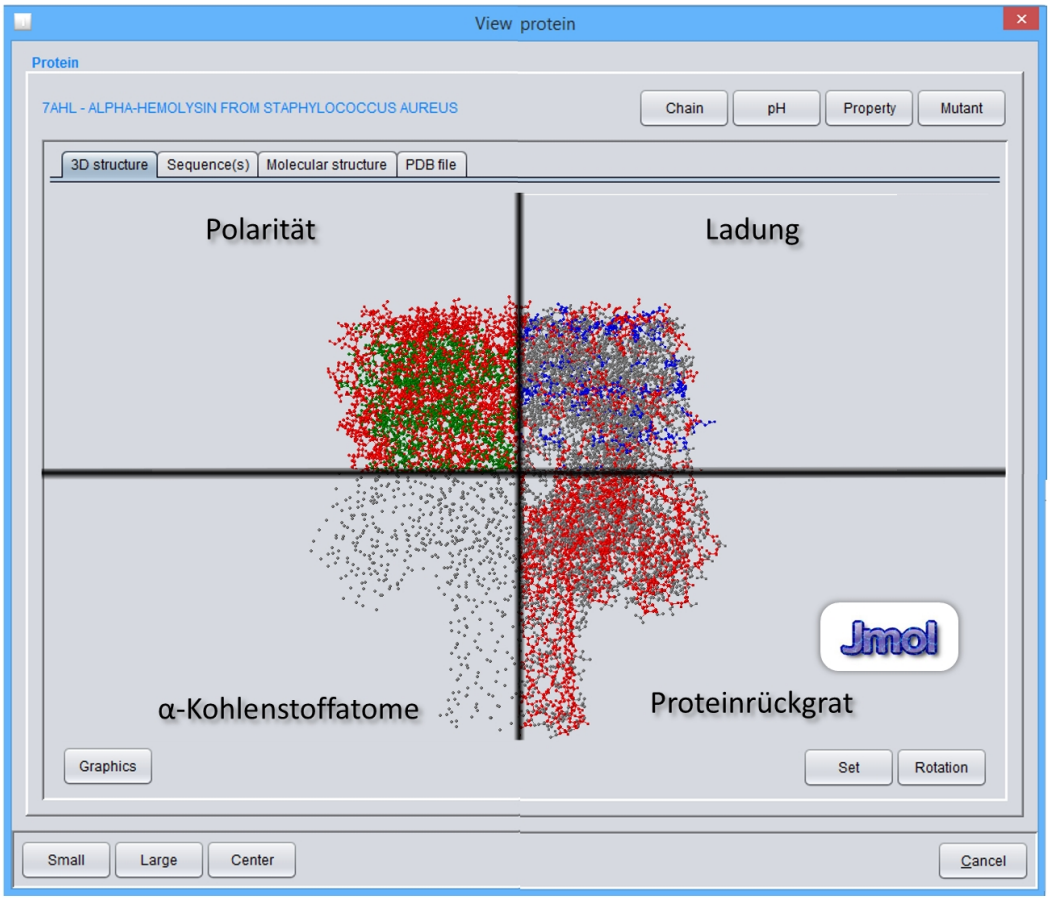


Abbildung 3.26: *Darstellung der vier möglichen Färbungen (Polarität, Ladung, α-Kohlenstoffatome und Proteinrückgrat) anhand des Proteins α-Hämolysin.*

Jmol unterstützt nur die Darstellung von chemischen Elementen und Bindungen. Für die Verwendung zur Darstellung von fragmentbasierten Simulati-

onsboxen muss ein Algorithmus entwickelt werden, welcher die unterschiedlichen Molekülfragmente in Abhängigkeit der gesetzten Farbe, Größe und Sichtbarkeit des jeweiligen Fragmentes auf die verschiedenen chemischen Elemente projiziert. Dabei musste das Problem ressourcensparend gelöst werden, da nur eine begrenzte Anzahl an chemischen Elementen zur Verfügung stehen. Abbildung 3.27 zeigt die Einbettung des `MfdJmolViewers` für die Darstellung von Simulationsboxen in MFD-FormulaOne. Zu sehen ist das in eine DMPC-Membran eingebettete Protein α-Hämolysin.

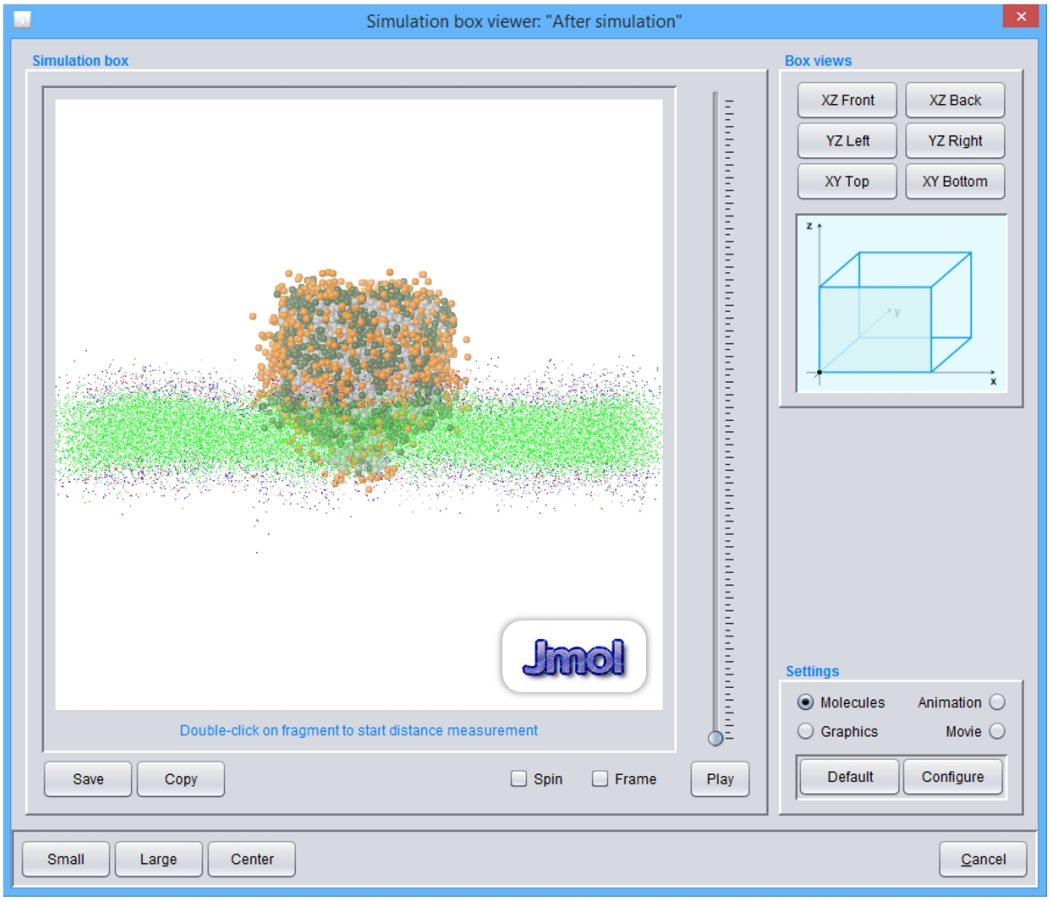


Abbildung 3.27: *Einbettung des MfdJmolViewers in die Anwendung MFD-FormulaOne zur Darstellung von Simulationsboxen. Zu sehen ist das Protein α-Hämolysin, orange: Hydrophiles Seitenkettenfragment, olivfarben: Hydrophobes Seitenkettenfragment, grau: Proteinrückgratfragment) von Staphylococcus aureus in einer DMPC-Membran (Grün: Hydrophobe Schicht, rot/blau: Positives/negatives Kopfgruppenfragment)*

3.2.4 Die Anwendung MfdCompassFX

Die Anwendung `MfdCompassFX` dient speziell der Berechnung des DPD-Repulsionsparameters und der hierfür benötigten Kenngrößen. Sie wird in Java 1.8 auf Basis des JavaFX Frameworks erstellt und umfasst 5188 Codezeilen. Das Programm bietet eine umfassende Nutzeroberfläche, welche sich in zwei Bereiche aufteilt: (1) Bestimmung von Moleküleigenschaften sowie der paarweisen Wechselwirkungsenergien und (2) Berechnung des DPD-Repulsionsparameters anhand der aus (1) gewonnen Daten. Für die Kalkulation der paarweisen Wechselwirkungsenergien erstellt `MfdCompass-FX` alle nötigen Skripte und Molekülgeometrien für die direkte Anwendung im Softwarepaket Discover 96.0 / 4.0.0 der Firma MSI [103] vollkommen automatisiert. Die berechneten Daten können anschließend zur Berechnung des DPD-Repulsionsparameters sowie aller nötigen Werte für das zu berechnende Fragmentset ausgewertet und an die Anwendung `MfdFragmentManager` (siehe Kapitel 3.2.5) weitergeleitet werden. Alle Molekuloperationen sowie die Symmetriebestimmung von Schweratomen basieren auf Algorithmen des *Chemistry Development Kit* (CDK) [166,167].

Die verwendeten Moleküle in `MfdCompassFX` müssen in Form des *InsightII* Dateiformats (Dateiendung: *.mfd/.cor*) vorliegen. Alle berechneten und erstellten Daten werden als Tabellen im Form von komma-separierten Werten (*Comma-separated values*, Dateiendung: *.csv*) organisiert und abgespeichert. Die für MFD-Simulationen nötigen Kenngrößen wie Molekülmasse und die lösungsmittelzugängliche Oberfläche (engl. *Solvent-Accessible Surface Area*) mit einer Sondengröße von 1.4 Å sind mit `MfdCompassFX` mit nur wenigen Mausklicks bestimmbar.

Zwei unterschiedliche Ansätze zur Bestimmung der paarweisen Wechselwirkungsenergien werden unterstützt: Der Dimeransatz und die Boxmethode. Der Dimeransatz verfolgt die Strategie, die paarweisen Wechselwirkungsenergien anhand eines Molekülpaares zu bestimmen. Die möglichen Einstellungen für den Dimeransatz sind in Abbildung 3.28 dargestellt.

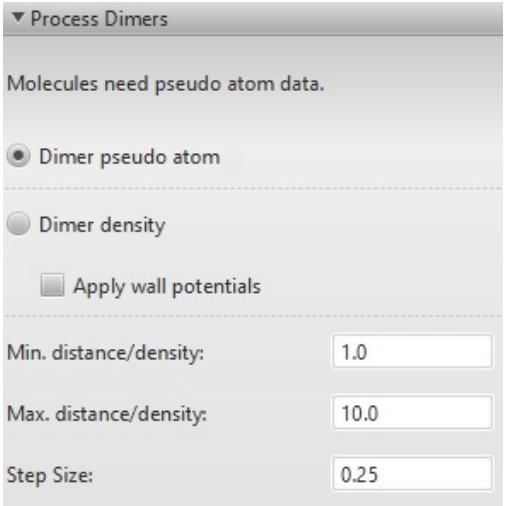


Abbildung 3.28: *Benutzeroberfläche zum Vornehmen von Einstellungen für die Ausführung des Dimeransatzes.*

Wenn der Pseudo-Atom-Dimeransatz (*Dimer pseudo atom*) gewählt ist, müssen alle Schweratome im Vorhinein mit `MfdCompassFX` bestimmt werden. Hierbei werden vorliegende Molekülsymmetrien berücksichtigt, sodass es nicht zu doppelten Berechnungen kommt. Für den dichteabhängigen Dimeransatz (*Dimere density*) wird zusätzlich die Option *Apply wall potentials* angeboten, um periodische Randbedingungen zu vermeiden. Für beide Ansätze sind die jeweilige minimale und maximale Distanz bzw. Dichte sowie die Schrittweite zu definieren.

Für die Verwendung des Boxansatzes muss zunächst die Anzahl der Moleküle in der Simulationsbox (*# Molecules*), die Anzahl der Wiederholungen pro Simulation (*# Samples*) und die vorherrschende Dichte in der Simulationsbox (*Density*) festgelegt werden. Die Einstellung *Distance-dependent dielectric* (deutsch Distanzabhängiges Dielektrikum) hat zur Folge, dass die Dielektrizitätskonstante als distanzabhängige Funktion beschrieben wird. Dies ist ein pragmatischer Ansatz zur genaueren Beschreibung von elektrostatischen Wechselwirkungen, wie sie in wässrigen Lösungen vorherrschen[173]. Die Option *Apply wall potentials* gleicht derjenigen des Dimeransatzes und unterbindet die Verwendung von periodischen Randbedingungen. Die Konfiguration *van der Waals (vdW) tail correction* (dt. vdW-Endenkorrektur) lässt bei Aktivierung die vdW-Kraft geglättet gegen null laufen (Abbildung 3.29), wodurch keine Diskontinuität im Verlauf der Kraft auftritt.

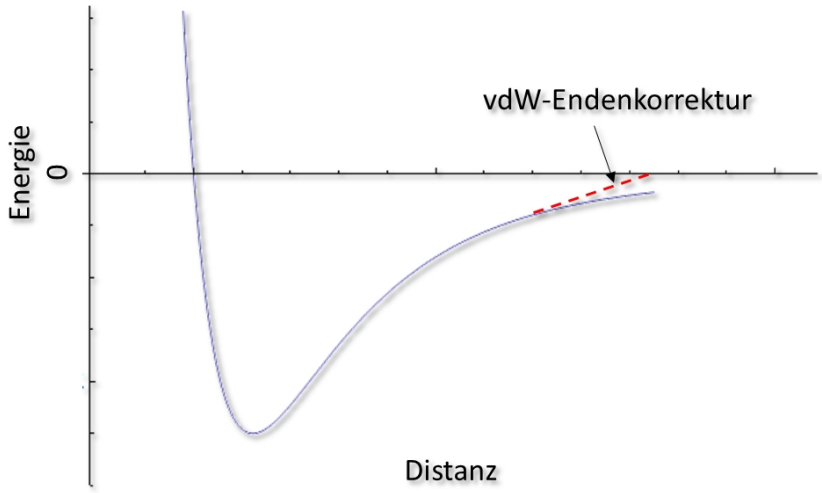


Abbildung 3.29: *Skizzierung des Verlaufs der vdW-Kraft bei aktivierter vdW-Endenkorrektur.*

Zur Verfügung stehen zwei Boxmethoden: Die erste Methode (*Single center*) erlaubt die Berechnung der paarweisen Wechselwirkungsenergien aus einer Box mit einem zentral platzierten Molekül umgeben von $N-1$ Molekülen des zweiten Typs. Der zweite Ansatz nutzt eine gleich verteilte Simulationsbox mit gleichen Anteilen beider Molekültypen. Abbildung 3.30 zeigt die grafische Benutzeropberfläche des Boxansatzes.

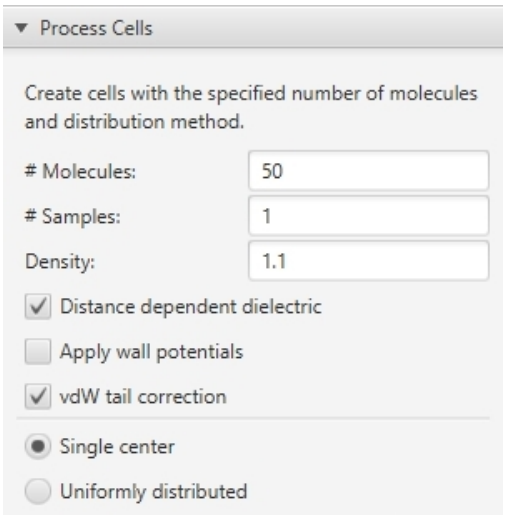


Abbildung 3.30: *Benutzeroberfläche zum Vornehmen von Einstellungen für die Ausführung der Boxmethode.*

Für jede Berechnungsmethode müssen die Temperatur in Kelvin und die Anzahl der Heizschritte sowie die Anzahl der Schritte für die Messung der Berechnung der paarweisen Wechselwirkungsenergie angegeben werden.

Für die Verarbeitung und Aufbereitung der berechneten paarweisen Wechselwirkungsenergien stehen vier Funktionen zur Verfügung (siehe Abbildung 3.31): Die Auswahl der Methode *Create fragment set* dient der Berechnung der DPD-Repulsionsparameter aus den paarweisen Wechselwirkungsenergien. Bei Angabe der Koordinationszahlen für die einzelnen Molekülpaare werden die paarweisen Wechselwirkungsenergien anhand von Gleichung 2.26 berechnet, ansonsten wird die Koordinationszahl als $Z_{ij} = 1$ angenommen. Im Anschluss an die Berechnung wird automatisch die Anwendung `MfdFragmentManager` zur Betrachtung der Ergebnisse geöffnet.

Die Funktion *Process raw energy data* eliminiert die intramolekularen nichtbindenden Wechselwirkungen bei Molekülen mit mehr als drei Atomen:

$$E_{ij} = E_{ij,raw} - (E_{i,g} + E_{j,g}) \tag{3.10}$$

wobei $E_{ij,raw}$ die paarweise Wechselwirkungsenergie inklusive intramolekularer Wechselwirkungen und $E_{i,g}$ die reine intramolekulare Wechselwirkungsenergie eines Moleküls in der Gasphase sind.

Die Methode *Reprocess dimere cell data* ermöglicht die Berechnung der Energieminima aus dichteabhängigen Dimeransatzberechnungen. Hierbei kann das zu betrachtende Dichteintervall neu definiert werden. Wenn die Eigenschaft *Create plots* aktiviert ist, wird zu jeder Berechnung ein Diagramm des Energieverlaufs in Abhängigkeit der Dichte gespeichert.

Die letzte Auswahlmöglichkeit (*Replace data with*) erlaubt dem Nutzer die Daten des ursprünglichen Datensets mit der Angabe von zusätzlichen Energiedaten zu überschreiben.

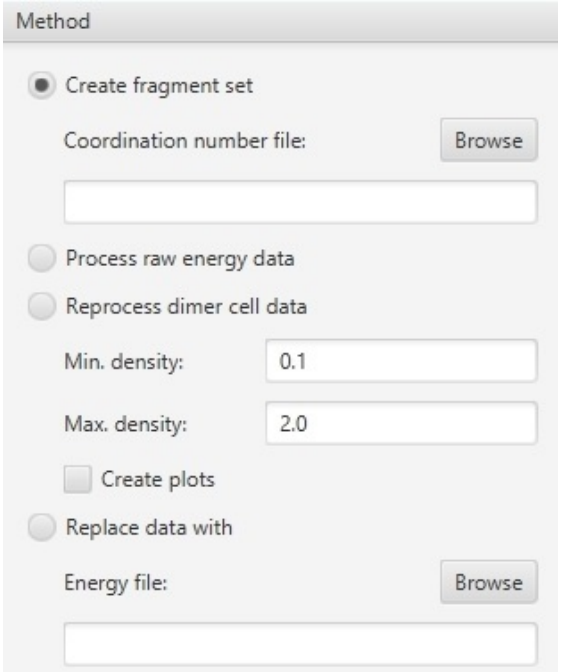
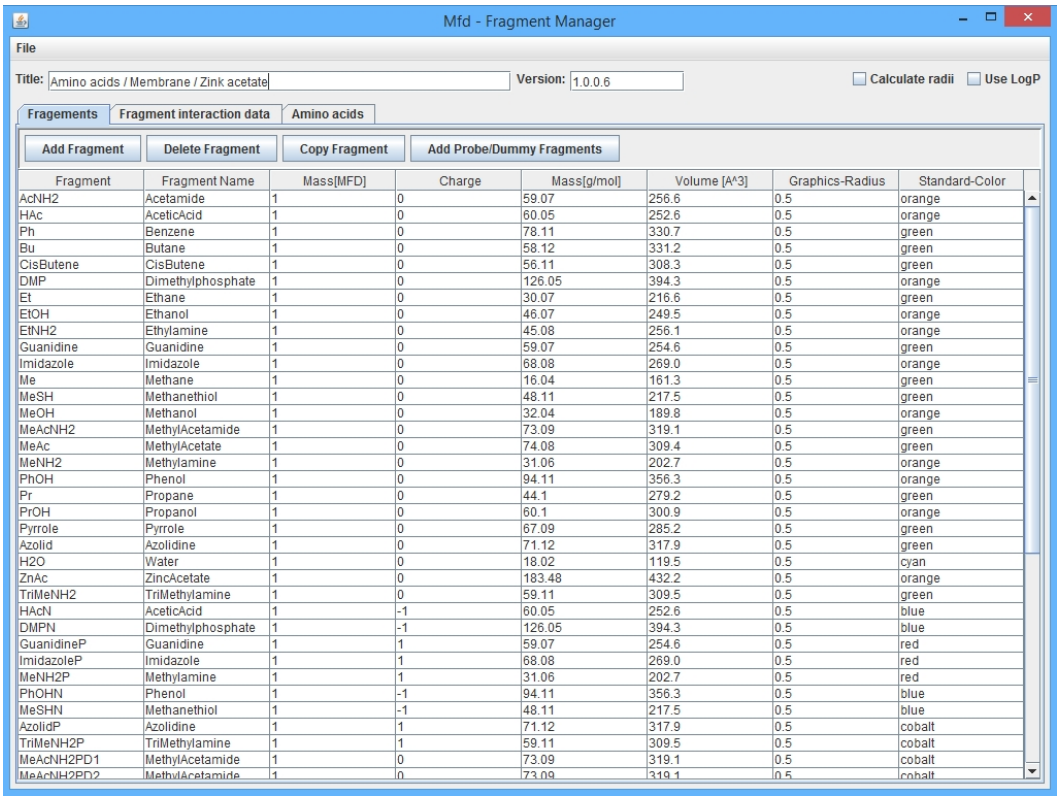


Abbildung 3.31: *Benutzeroberfläche zum Prozessieren der mit MfdCompassFX berechneten Daten.*

3.2.5 Die Anwendung MfdFragmentManager

Die Anwendung `MfdFragmentManager` bietet eine grafische Oberfläche für die Verwaltung von Molekülfragmenten und deren Eigenschaften, den zugehörigen DPD-Repulsionsparametern a_{ij} und die Struktur von den aus Molekülfragmenten aufgebauten Aminosäuren. Jeder Aufgabenbereich ist in ein eigenes sogenanntes *Tab* organisiert: *Fragments*, *Fragment interaction data* und *Amino acids*. Abbildung 3.32 zeigt den Startbildschirm der Applikation mit dem geöffneten Standardfragmentset für Proteine.



Fragment	Fragment Name	Mass[MFD]	Charge	Mass[g/mol]	Volume [A^3]	Graphics-Radius	Standard-Color
AcNH2	Acetamide	1	0	59.07	256.6	0.5	orange
HAc	AceticAcid	1	0	60.05	252.6	0.5	orange
Ph	Benzene	1	0	78.11	330.7	0.5	green
Bu	Butane	1	0	58.12	331.2	0.5	green
CisButene	CisButene	1	0	56.11	308.3	0.5	green
DMP	Dimethylphosphate	1	0	126.05	394.3	0.5	orange
Et	Ethane	1	0	30.07	216.6	0.5	green
EtOH	Ethanol	1	0	46.07	249.5	0.5	orange
EtNH2	Ethylamine	1	0	45.08	256.1	0.5	orange
Guanidine	Guanidine	1	0	59.07	254.6	0.5	green
Imidazole	Imidazole	1	0	68.08	269.0	0.5	orange
Me	Methane	1	0	16.04	161.3	0.5	green
MeSH	Methanethiol	1	0	48.11	217.5	0.5	green
MeOH	Methanol	1	0	32.04	189.8	0.5	orange
MeAcNH2	MethylAcetamide	1	0	73.09	319.1	0.5	green
MeAc	MethylAcetate	1	0	74.08	309.4	0.5	green
MeNH2	Methylamine	1	0	31.06	202.7	0.5	orange
PhOH	Phenol	1	0	94.11	356.3	0.5	orange
Pr	Propane	1	0	44.1	279.2	0.5	green
PrOH	Propanol	1	0	60.1	300.9	0.5	orange
Pyrrole	Pyrrole	1	0	67.09	285.2	0.5	green
Azolid	Azolidine	1	0	71.12	317.9	0.5	green
H2O	Water	1	0	18.02	119.5	0.5	cyan
ZnAc	ZincAcetate	1	0	183.48	432.2	0.5	orange
TriMeNH2	TriMethylamine	1	0	59.11	309.5	0.5	green
HAcN	AceticAcid	1	-1	60.05	252.6	0.5	blue
DMPN	Dimethylphosphate	1	-1	126.05	394.3	0.5	blue
GuanidineP	Guanidine	1	1	59.07	254.6	0.5	red
ImidazoleP	Imidazole	1	1	68.08	269.0	0.5	red
MeNH2P	Methylamine	1	1	31.06	202.7	0.5	red
PhOHN	Phenol	1	-1	94.11	356.3	0.5	blue
MeSHN	Methanethiol	1	-1	48.11	217.5	0.5	blue
AzolidP	Azolidine	1	1	71.12	317.9	0.5	cobalt
TriMeNH2P	TriMethylamine	1	1	59.11	309.5	0.5	cobalt
MeAcNH2PD1	MethylAcetamide	1	0	73.09	319.1	0.5	cobalt
MeAcNH2PD2	MethylAcetamide	1	0	73.09	319.1	0.5	cobalt

Abbildung 3.32: *Startbildschirm mit geöffnetem Standardfragmentset für Proteine.*

Das Hauptfenster *Fragments* ermöglicht die Verwaltung der verfügbaren Fragmente. Neben dem Titel und der Version sind folgende Spalten verfügbar und editierbar:

- Fragment: Abkürzung des Fragmentes.
- FragmentName: Chemischer Name des Fragmentes.
- Mass[MFD]: Masse des Fragmentes in MFD-Einheiten (immer 1).
- Charge: Ladung des Fragments.
- Mass[g/mol]: Masse des Fragments in $g \cdot mol^{-1}$.
- Volume[Å^3]: Volumen des Fragmentes. Abgeleitet von der lösungsmittelzugänglichen Oberfläche für eine 1.4 Å große Sonde.
- Graphics-Radius: Radius des Fragmentes in Ångstrom (nicht verwendet).
- Standard-Color: Anzeigefarbe des Fragmentes in Mfd-FormulaOne.

Zusätzlich gibt es die Möglichkeit, Fragmente manuell zu erstellen (*Add Fragment* Schaltfläche) und Fragmente zu löschen (*Delete Fragment* Schaltfläche). Die Schaltfläche *Copy Fragment* bietet dem Anwender an, ein gegebenes Fragment zu kopieren. Soweit *Apply Ionic Shift* nicht aktiviert ist, werden die ursprünglichen a_{ij}-Parameter dupliziert. Mit der Aktivierung von *Apply Ionic Shift* werden die Regeln für geladene Fragmente (siehe Kapitel 2.1.5) angewendet. Dabei wird als Referenzsubstanz die Wahl unter *Reference* verwendet und α wie unter dem gleichnamigen Menüpunkt *Alpha* gewählt. Der in Abbildung 3.33 gezeigte Dialog produziert bei Bestätigung ein negativ geladenes Essigsäurefragment mit der Formalladung -1, mit Wasser (H_2O) als Referenzsubstanz und $\alpha = 1.0$.

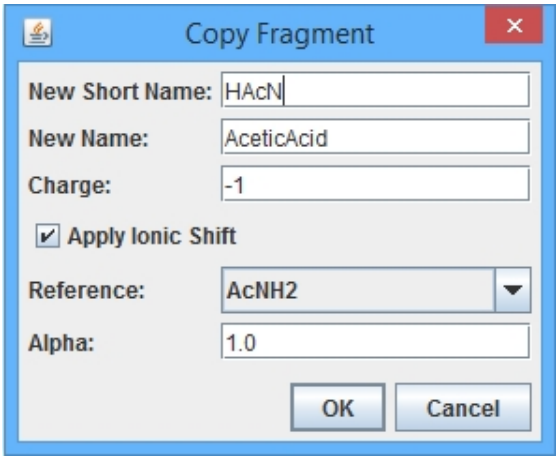


Abbildung 3.33: *Dialog zum kopieren von Fragmenten.*

Im Dialog *Add Probe/Dummy Fragment* lassen sich Fragmente n-mal duplizieren und bekommen die Endung *PDX*, wobei *X* eine fortlaufende Nummer ist. Diese so erstellten Sonden- bzw. Dummyfragmente können innerhalb des Proteinrückgrates gesetzt werden, um z.B. die Geometrie eines Proteins während einer Simulation zu analysieren.

Der zweite Bildschirm des `MfdFragmentManagers` dient der Visualisierung und Manipulation der a_{ij}-Parameter. Die Schaltflächen besitzen folgende Funktionalität:

- *Add Temperature*:

Ermöglicht das Hinzufügen eines zusätzlichen Fragmentsets einer noch nicht vorhandenen Temperatur. Die Fragmente des einzufügenden Sets müssen hierbei deckungsgleich mit dem aktuellen Set sein.

- *Change Temperature*:

  Dient der Änderung des Zahlenwertes der fokussierten Temperatur.

- *Delete Temperature*:

  Löscht das aktuell fokussierte Fragmentset.

- *Normalize*:

  Skalierung der a_{ij}-Parameter, sodass die Spannweite der Parameter einem definierten Minimum $a_{ij,newmin}$ und Maximum $a_{ij,newmax}$ entsprechen:

$$a_{ij,new} = \frac{(a_{ij,old} - a_{ij,oldmin}) \cdot (a_{ij,newmax} - a_{ij,newmin})}{a_{ij,oldmax} - a_{ij,oldmin}} + a_{ij,newmin} \tag{3.11}$$

  wobei $a_{ij,oldmin}$ das momentane Minimum und $a_{ij,oldmax}$ das momentane Maximum der fokussierten a_{ij}-Parameter und $a_{ij,old}$ der fokussierte unskalierte a_{ij}-Parameter ist.

- *Clean*:

  Entfernt alle Fragmente, die ungültige a_{ij}-Parameter besitzen, d.h. die Methode generiert ein lauffähiges Minimalfragmentset.

- *Constrain IA*:

  Öffnet den Dialog *Constrain Interactions*, welcher dem Benutzer ermöglicht, Grenzwerte der a_{ij}-Parameter zu definieren. Das Textfeld *Min. uncharged* und *Max. uncharged* definieren jeweils den unteren bzw. oberen Grenzwert. Initialisiert sind die Textfelder mit dem entsprechenden minimalen bzw. maximalen a_{ij}-Parameter der ungeladenen Fragmente. Mittels der letzten drei Textfelder lassen sich die Grenzwerte, definiert durch die ersten beiden Werte, aufweiten, sodass in Einheiten der Spannweite der angegeben Grenzen diese in beide Richtungen (Minimalwert und Maximalwert) erweitert werden können. Somit ist die Einstellung der

Grenzen für Wechselwirkungen von Fragmenten die ungeladen, ungleich geladen und gleich geladen sind, gesondert möglich.

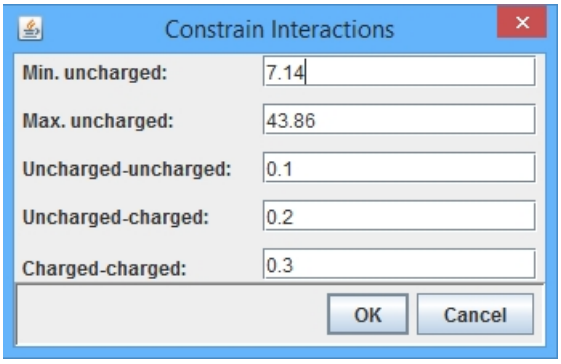


Abbildung 3.34: *Dialog zum Definieren von Grenzwerten der a_{ij}-Parameter.*

- *Refine Charged IA*:

  Wendet die Gleichungen 3.5 und 3.6 auf die a_{ij}-Parameter der geladenen Fragmente an, sodass die Wechselwirkungen der gleich geladenen Fragmente die größte Repulsion bzw. die Wechselwirkungen der ungleich geladenen Fragmente die höchste Attraktion aller Parameter im Temperaturset besitzen. Der Korrekturfaktor $c_{correct}$ kann über einen Dialog definiert werden.

Unterhalb der Schaltflächen findet sich ein sogenannter *Slider*. Dieser dient der Feinabstimmung des Diagonalelements der Repulsionsparameter. Das Diagonalelement bestimmt die Grenze zu repulsiver und attraktiver Wechselwirkung. Die *Checkbox Link Temperatures* führt dazu, wenn aktiviert, dass sich Änderungen auf alle Temperaturen gleichzeitig auswirken und die einzelnen Fragmentsets der verschiedenen Temperaturen nicht getrennt voneinander betrachtet werden.

Die untere Tabelle in Abbildung 3.35 zeigt eine grafische Darstellung der Repulsionsparameter und ist organisiert in Form einer Dreiecksmatrix.

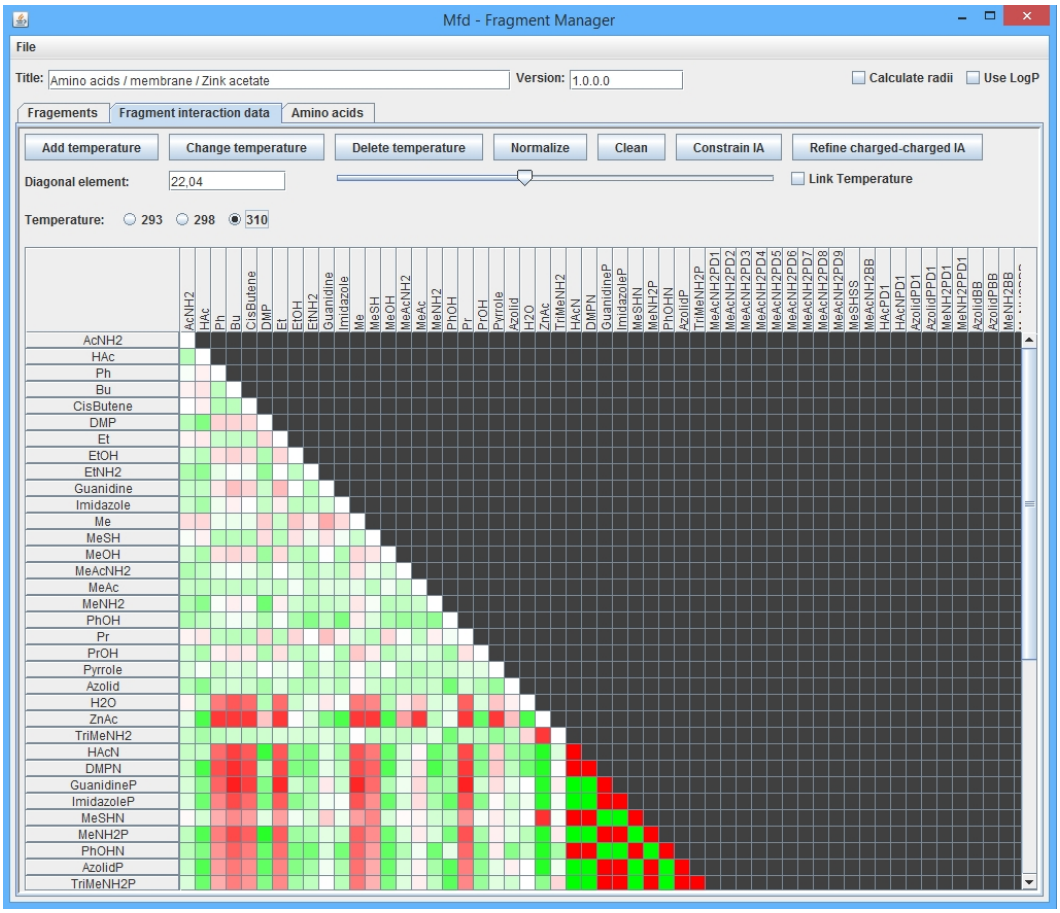


Abbildung 3.35: *Screenshot der Oberfläche zur Visualisierung und Bearbeitung der a_{ij}-Parameter mit geöffnetem Standardfragmentset für Proteine.*

Abbildung 3.36 zeigt die zugehörige Farbskalierung der Repulsionsparameter. Neutrale Werte sind weiß dargestellt. Je intensiver grün bzw. rot die entsprechende Zelle des Repulsionsparameters ist, desto attraktiver bzw. repulsiver verhalten sich die entsprechenden Fragmente in der Simulation untereinander.

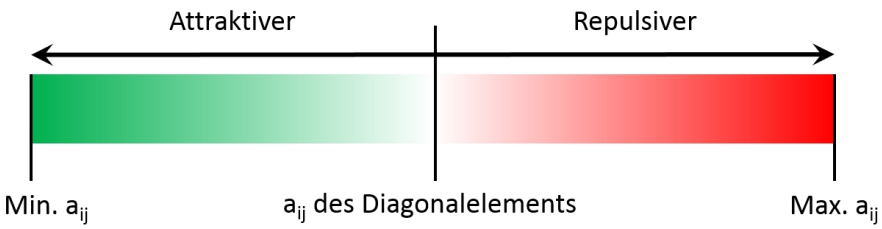
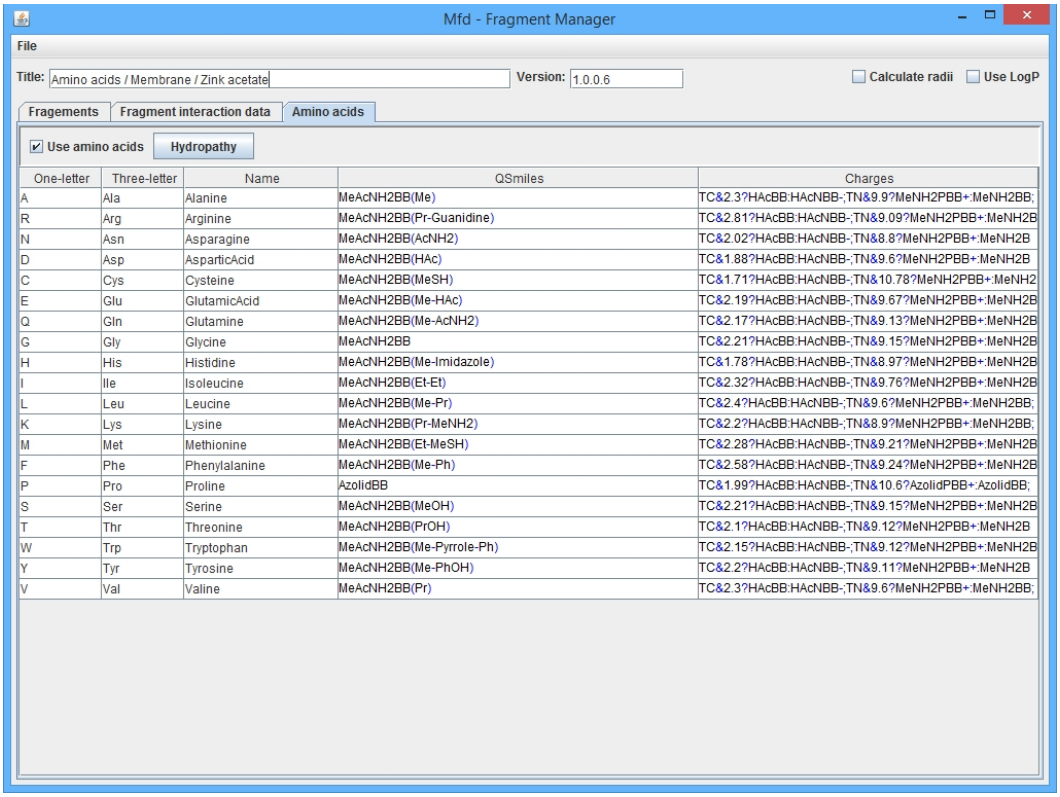


Abbildung 3.36: *Farbskala der a_{ij}-Repulsionparameter.*

Das letzte Fenster *AminoAcids* bietet eine Tabelle, welche alle nötigen Informationen über Aminosäuren bezüglich MFD-Simulationen enthält. Die Bereitstellung der Daten ist nicht zwingend und kann mittels des Kontrollkästchens *Use amino acids* aktiviert bzw. deaktiviert werden. Die Tabelle beinhaltet den Ein- bzw. Dreibuchstabencode, den Namen, den *f*SMILES und die Kodierung der pKs-Werte der einzelnen Aminosäuren. Die gesamte in dieser Arbeit verwendete Tabelle mit den Daten aller Aminosäuren ist im Anhang 11.2 zu finden.



One-letter	Three-letter	Name	QSmiles	Charges
A	Ala	Alanine	MeAcNH2BB(Me)	TC&2.3?HAcBB:HAcNBB-;TN&9.9?MeNH2PBB+:MeNH2BB;
R	Arg	Arginine	MeAcNH2BB(Pr-Guanidine)	TC&2.81?HAcBB:HAcNBB-;TN&9.09?MeNH2PBB+:MeNH2B
N	Asn	Asparagine	MeAcNH2BB(AcNH2)	TC&2.02?HAcBB:HAcNBB-;TN&8.8?MeNH2PBB+:MeNH2B
D	Asp	AsparticAcid	MeAcNH2BB(HAc)	TC&1.88?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2B
C	Cys	Cysteine	MeAcNH2BB(MeSH)	TC&1.71?HAcBB:HAcNBB-;TN&10.78?MeNH2PBB+:MeNH2
E	Glu	GlutamicAcid	MeAcNH2BB(Me-HAc)	TC&2.19?HAcBB:HAcNBB-;TN&9.67?MeNH2PBB+:MeNH2B
Q	Gln	Glutamine	MeAcNH2BB(Me-AcNH2)	TC&2.17?HAcBB:HAcNBB-;TN&9.13?MeNH2PBB+:MeNH2B
G	Gly	Glycine	MeAcNH2BB	TC&2.21?HAcBB:HAcNBB-;TN&9.15?MeNH2PBB+:MeNH2B
H	His	Histidine	MeAcNH2BB(Me-Imidazole)	TC&1.78?HAcBB:HAcNBB-;TN&8.97?MeNH2PBB+:MeNH2B
I	Ile	Isoleucine	MeAcNH2BB(Et-Et)	TC&2.32?HAcBB:HAcNBB-;TN&9.76?MeNH2PBB+:MeNH2B
L	Leu	Leucine	MeAcNH2BB(Me-Pr)	TC&2.4?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2BB;
K	Lys	Lysine	MeAcNH2BB(Pr-MeNH2)	TC&2.2?HAcBB:HAcNBB-;TN&8.9?MeNH2PBB+:MeNH2BB;
M	Met	Methionine	MeAcNH2BB(Et-MeSH)	TC&2.28?HAcBB:HAcNBB-;TN&9.21?MeNH2PBB+:MeNH2B
F	Phe	Phenylalanine	MeAcNH2BB(Me-Ph)	TC&2.58?HAcBB:HAcNBB-;TN&9.24?MeNH2PBB+:MeNH2B
P	Pro	Proline	AzolidBB	TC&1.99?HAcBB:HAcNBB-;TN&10.6?AzolidPBB+:AzolidBB;
S	Ser	Serine	MeAcNH2BB(MeOH)	TC&2.21?HAcBB:HAcNBB-;TN&9.15?MeNH2PBB+:MeNH2B
T	Thr	Threonine	MeAcNH2BB(PrOH)	TC&2.1?HAcBB:HAcNBB-;TN&9.12?MeNH2PBB+:MeNH2B
W	Trp	Tryptophan	MeAcNH2BB(Me-Pyrrole-Ph)	TC&2.15?HAcBB:HAcNBB-;TN&9.12?MeNH2PBB+:MeNH2B
Y	Tyr	Tyrosine	MeAcNH2BB(Me-PhOH)	TC&2.2?HAcBB:HAcNBB-;TN&9.11?MeNH2PBB+:MeNH2B
V	Val	Valine	MeAcNH2BB(Pr)	TC&2.3?HAcBB:HAcNBB-;TN&9.6?MeNH2PBB+:MeNH2BB;

Abbildung 3.37: *Screenshot des Fensters AminoAcids.*

4 Vorstudie von nicht-ionischen Tensiden an einer Wasser-Luft-Grenzfläche

Für die Validierung der MFD-Methode wird eine Vorstudie durchgeführt, die das Verhalten von nicht-ionischen Polyoxyethylenalkylethern an einer Wasser-Luft-Grenzfläche zum Gegenstand hat[48]. Betrachtet wird im Folgenden die Bildung von nanoskaligen Strukturen und die Grenzflächeneigenschaften der Tenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$. Die Ergebnisse dieser Untersuchung werden anhand von experimentellen Daten validiert. Abbildung 4.1 zeigt die molekulare Struktur der C_xE_y-Tenside.

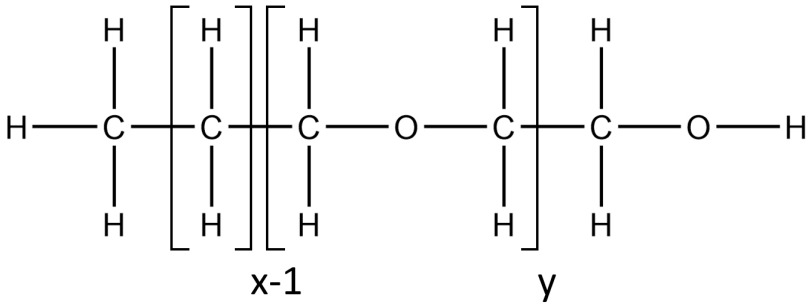


Abbildung 4.1: *Molekulare Struktur der Polyoxyethylenalkylether-Tenside C_xE_y.*

Methode

Für die Untersuchungen werden rechtwinkelige Simulationsboxen mit einer Größe von 30·30·45 nm^3 verwendet. Die Boxen enthalten $3.0 \cdot 10^5$ Wasser- (H_2O), $6.4 \cdot 10^4$ Stickstoff- (N_2) und $1.6 \cdot 10^4$ Sauerstofffragmente (Moleküle). Periodische Randbedingungen sind in x- und y-Richtung parallel zur Wasser-Luft-Grenzfläche aktiviert. Die Wasser und Stickstoff-/Sauerstofffragmente sind in zwei unterschiedlichen Schichten für die Ausbildung der Wasser-Luft-Grenzfläche organisiert (Abbildung 4.1a). Unter der Grenzfläche sind die Polyoxyethylenalkyletherf-Moleküle angeordnet (Abbildung 4.1b). Eine Simulation mit $2.0 \cdot 10^4$ Zeitschritten rechnet 10 h auf einem Kern einer Intel-Xeon Workstation.

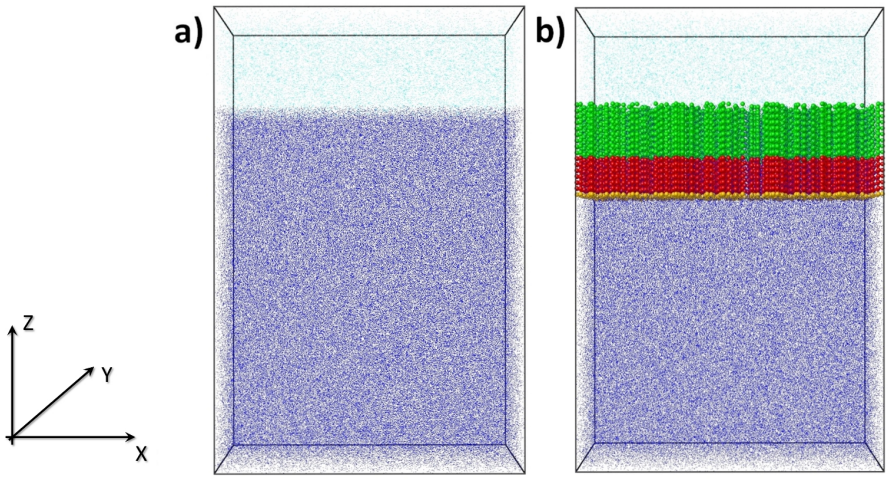


Abbildung 4.2: *a) Simulationsbox mit Wasser (blau) und Luft (Sauerstoff/Stickstoff) Segmenten (Cyan). b) Startgeometrie der Simulation mit unter der Wasser-Luft-Grenzfläche angeordneten Polyoxyethylenalkylether-Molekülen (Methan: Grün, Dimethylether: Rot, Methanol: Orange)*[48].

Um eine scharfe Trennung zwischen der Wasser-Luft-Grenzfläche aufrecht zu erhalten, werden die Repulsionsparameter a_{ij} der Fragmente Wasser, Dimethylether und Methanol zum Luftgemisch (Sauerstoff/Stickstoff) dahingehend angepasst, sodass sie stark repulsiv wirken. Das Fragmentierungsschema der C_xE_y Tenside ist in Abbildung 4.3 gezeigt. Jedes Tensidmolekül besteht aus $(x-1)$ Methanfragmenten, y Dimethyletherfragmenten und einem endständigen Metha-

nolfragment.

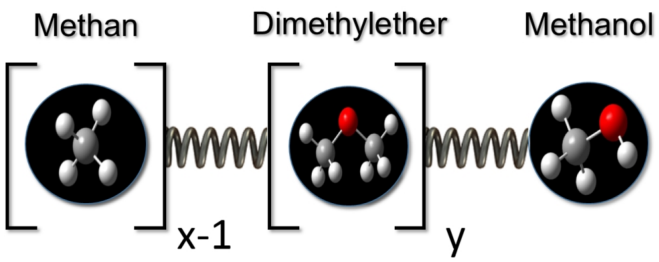


Abbildung 4.3: *Fragmentierungsschema der C_xE_y-Tenside*[48].

Ergebnis

Die Anzahl der Tensidmoleküle in der Simulationsbox variiert schrittweise von 100 bis 1500 Molekülen in einer Schrittweite von 100 Molekülen. Die berechnete Oberflächenspannung wird anhand von zwei Punkten skaliert. Der erste Punkt ist die experimentelle Oberflächenspannung einer reinen Wasser-Luft-Grenzfläche (γ_0 = 72.0 mN·m^{-1} bei 25 °C). Der zweite Punkt ist die kritische Mizellenkonzentration (englisch *critical micelle concentration*, kurz CMC) des Tensids C_6E_6 (γ_{CMC} = 32.0 mN·m^{-1} bei 25 °C[174], übereinstimmend mit 1000 C_6E_6 Molekülen in der Simulationsbox, siehe Abbildung 4.6).

Querschnitte der Simulationsbox mit nanoskaligen Strukturen im thermodynamischen Gleichgewicht des Tensids $C_{16}E_6$ in verschiedenen Konzentrationen sind in Abbildung 4.4 gezeigt. Die Bildung von Mizellen ist erkennbar bei der Konzentrationsänderung von 500 (keine Mizellen) auf 600 $C_{16}E_6$-Moleküle (Mizellen sichtbar).

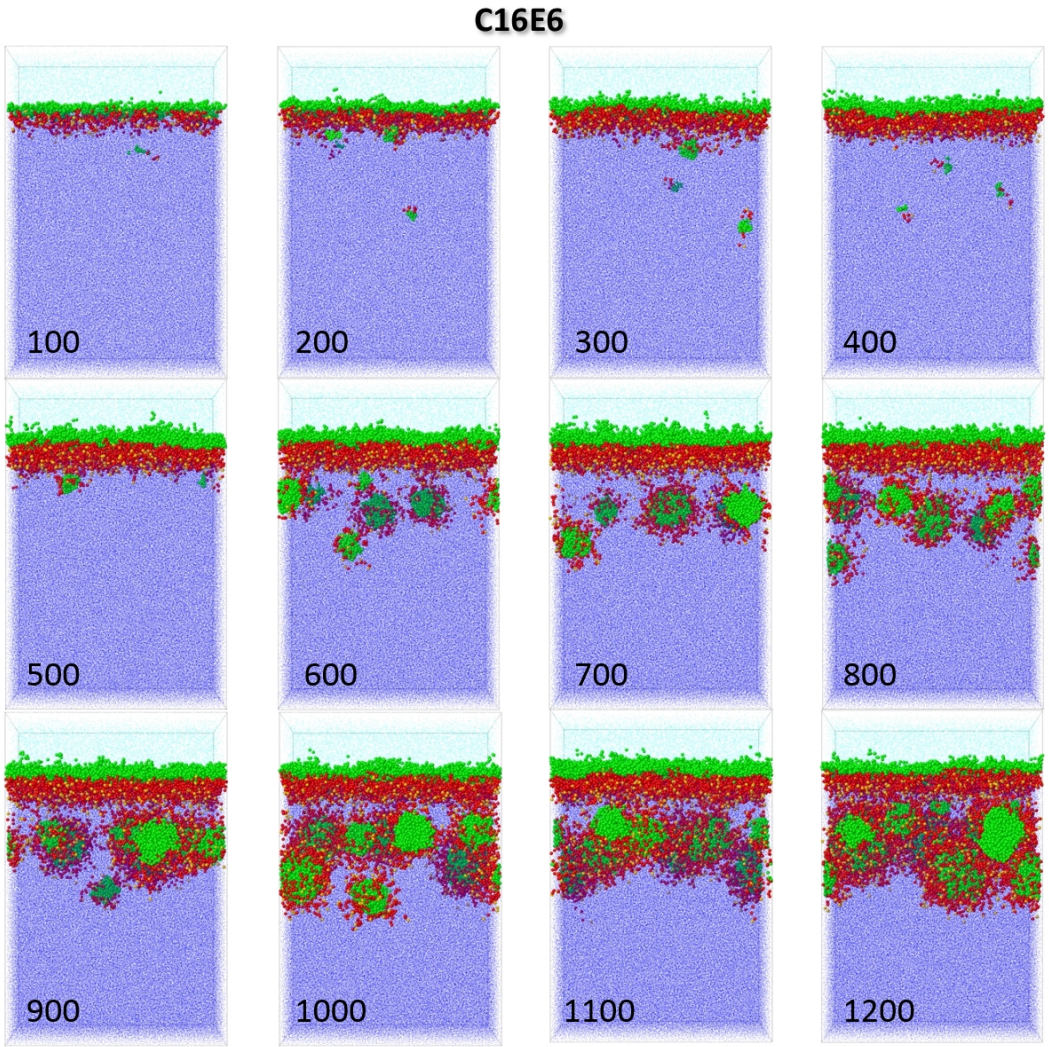


Abbildung 4.4: *Simulationsboxquerschnitte des Tensids $C_{16}E_6$ für die Konzentration von 100 bis 1200 Molekülen in Schritten von 100 Molekülen*[48].

In Abbildung 4.5 sind Querschnitte der Simulationsboxen aller untersuchten Tenside bei unterschiedlichen Tensidkonzentrationen zusammengefasst. Die Konzentration (Anzahl der Tensidmoleküle in der Box) in der mittleren Reihe zeigt, dass Tenside mit längeren C_x-Alkylketten eine niedrigere Konzentration aufweisen müssen, um Mizellen auszubilden. Dieses Ergebnis stimmt mit den experimentellen CMC-Konzentrationen überein, welche ebenso abnehmen, wenn die Länge der C_x-Alkylketten zunimmt[175].

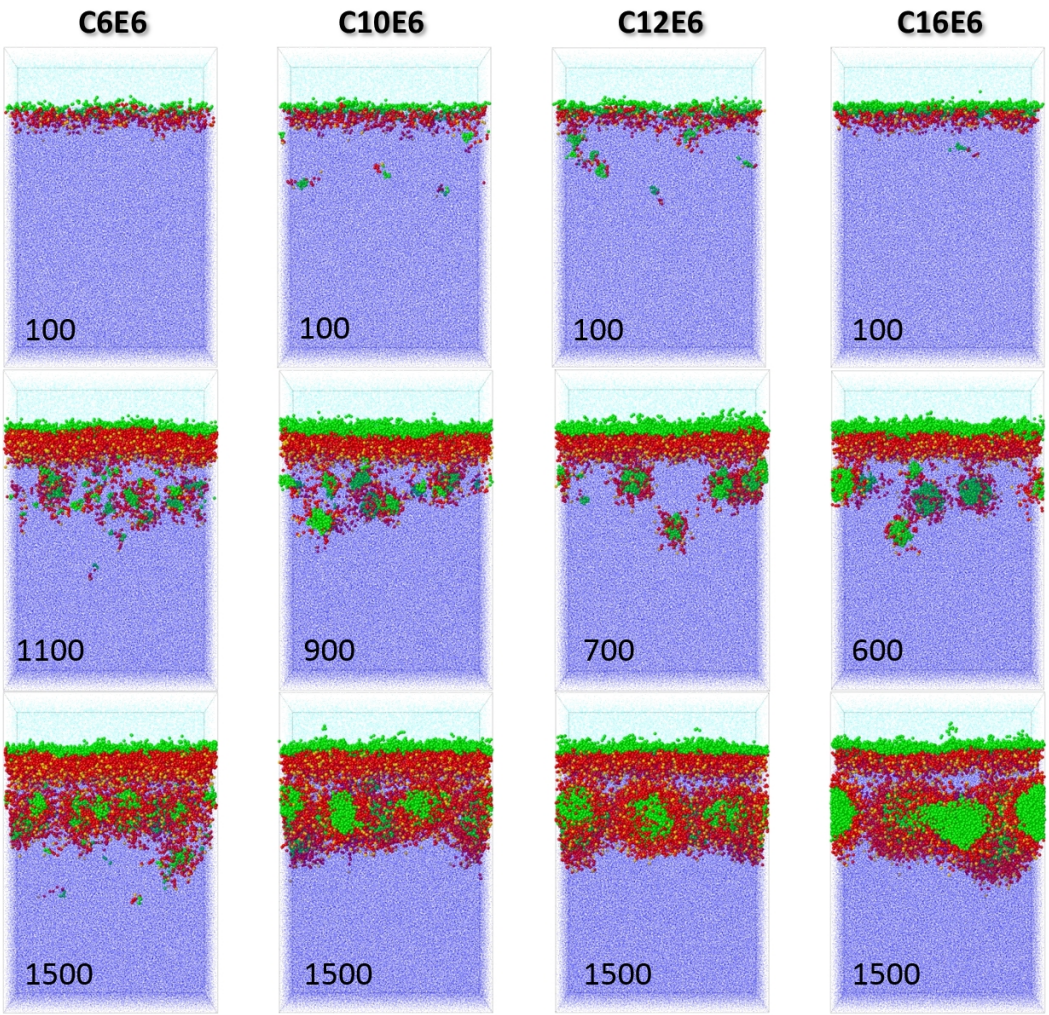


Abbildung 4.5: *Querschnitte der Simulationsboxen für verschiedene Konzentrationen der Tenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$ nachdem das thermodynamische Gleichgewicht erreicht ist. Die oberste und unterste Reihe zeigen jeweils die niedrigste bzw. höchste untersuchte Konzentration. Die mittlere Reihe zeigt die Konzentration, bei welcher die Ausbildung von Mizellen beginnt*[48].

Abbildung 4.6 fasst die berechneten Oberflächenspannungen bei verschiedenen Tensidkonzentrationen zusammen. Für eine Konzentration von beispielsweise 400 Tensidmolekülen in der Simulationsbox ist die Oberflächenspannung auf 58.07 $mN{\cdot}m^{-1}$ für C_6E_6, 54.25 $mN{\cdot}m^{-1}$ für $C_{10}E_6$, 51.42 $mN{\cdot}m^{-1}$ für $C_{12}E_6$ und 47.28 $mN{\cdot}m^{-1}$ für $C_{16}E_6$ herabgesetzt. Dies zeigt die erwartete Abschwächung gegenüber der reinen Wasser-Luft-Grenzfläche mit einem Wert von $\gamma_0 = 72.0$ $mN{\cdot}m^{-1}$. Eine längere C_x-Alkylkette führt zu einer geringeren berechneten Oberflächenspannung. Nachdem die Sättigung der Wasser-Luft-Grenzfläche mit Tensidmolekülen

erreicht ist, beginnt die Oberflächenspannung zu fluktuieren. Diese Fluktuationen können eventuell damit begründet werden, dass die Oberflächenspannung über die Aufsummierung der diagonalen Drucktensoren über die gesamte Box berechnet wird[176]. Dies führt dazu, dass die Beiträge der Wasser-Luft-Grenzfläche durch die Bildung der Mizellstrukturen verzerrt werden.

Die experimentelle Effektivität der Reduktion der Oberflächenspannung, Π_{CMC} $(= \gamma_0 - \gamma_{CMC})$, wobei γ_0 die Oberflächenspannung der Wasser-Luft-Grenzfläche und γ_{CMC} die Oberflächenspannung bei CMC ist, ist die Effektivität ungefähr $\Pi_{CMC} = 40$ mN·m^{-1} für alle beobachteten Tenside[174]. Die berechneten Werte zeigen eine abfallende Effektivität der Reduktion der Oberflächenspannung für eine steigende C_x Alkylkettenlänge: 40 mN·m^{-1} für C_6E_6, 37 für $C_{10}E_6$, 31 für $C_{12}E_6$ und 30 für $C_{16}E_6$.

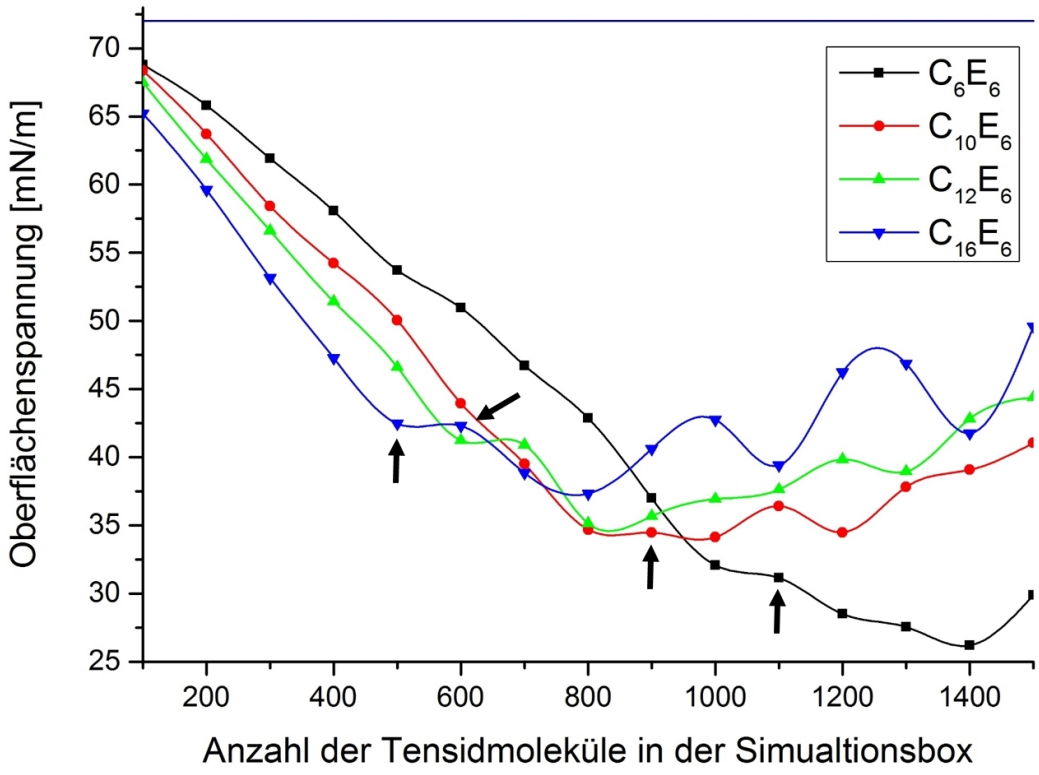


Abbildung 4.6: *Berechnete Oberflächenspannung für verschiedene Konzentrationen der Tenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$. Die Pfeile kennzeichnen die erste Bildung von Mizellen*[48].

Die Effektivität der Tensidabsorption an einer Oberfläche ist definiert als der

negative Logarithmus der Konzentration des Tensids, welche benötigt wird, um eine Reduktion der Oberflächenspannung des Lösungsmittels von 20 mN·m^{-1} zu produzieren, $pC_{20} \equiv -\log C_{(-\Delta\gamma=20)}$. Somit führt eine Erhöhung des pC_{20}-Wertes zu einer reziproken Abschwächung der Oberflächenspannung. Abbildung 4.7 zeigt den Vergleich der experimentellen pC_{20}-Werte[174] mit den korrespondierenden berechneten Oberflächenspannungen in Abhängigkeit der C_x-Alkylkettenlänge, um die Übereinstimmung zu demonstrieren.

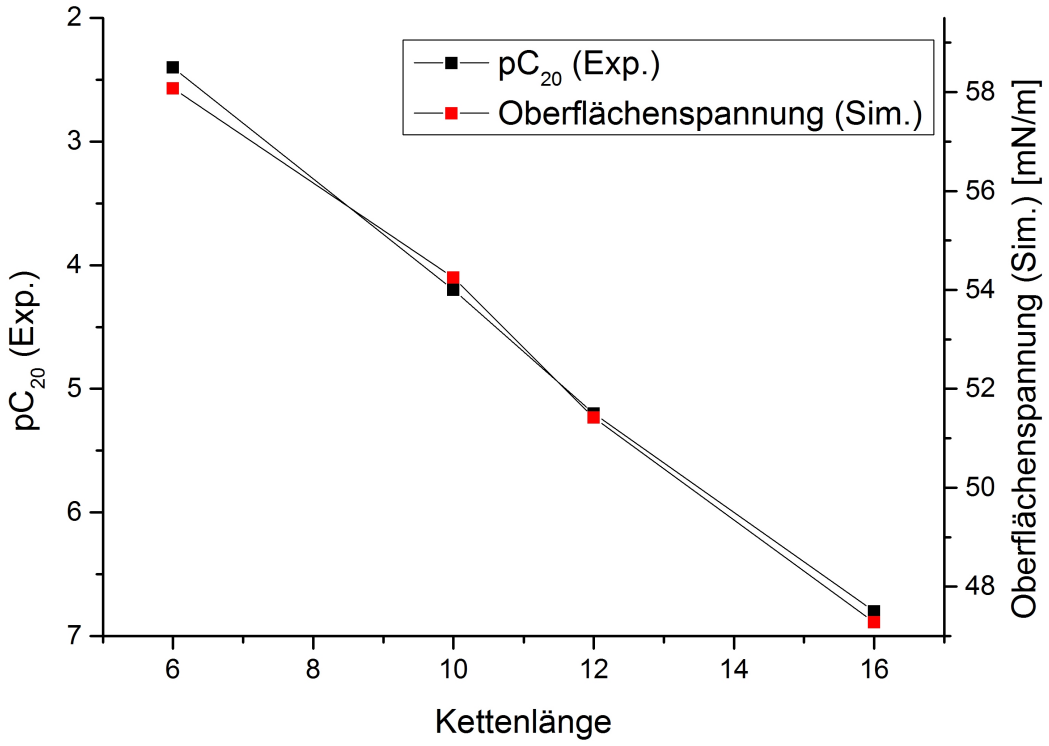


Abbildung 4.7: *Experimentelle pC_{20}-Werte[174] und die berechneten Oberflächenspannungen für eine Tensidkonzentration von 400 Molekülen in Abhängigkeit von der C_x-Alkylkettenlänge der Tenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$[48].*

Die Anzahl der Tensidmoleküle an der Wasser-Luft-Grenzfläche pro nm^2 gegenüber der Gesamtanzahl an Tensidmolekülen in der Simulationsbox ist in Abbildung 4.8 dargestellt. Die Anzahl der Tensidmoleküle steigt linear mit der Erhöhung der Konzentration bis zur Oberflächensättigung an. Dieses Ergebnis ist direkt vergleichbar mit der Abhängigkeit der berechneten Oberflächenspannung der Tensidkonzentration (siehe Abbildung 4.6).

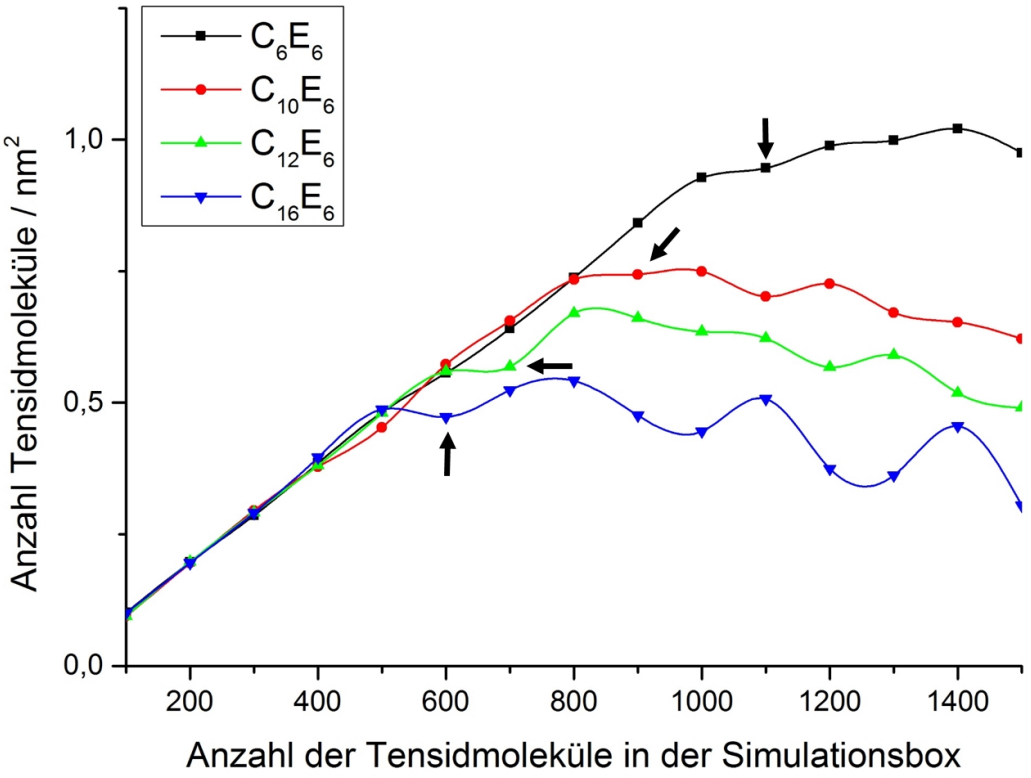


Abbildung 4.8: *Anzahl der Tensidmoleküle an der Wasser-Luft-Grenzfläche in Abhängigkeit der Gesamtanzahl von Tensidmolekülen in der Simulationsbox für die Tenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$. Die Pfeile kennzeichnen die erste Bildung von Mizellen*[48].

Die durchschnittliche Dicke der von hydrophoben (grün) Fragmenten geformten Schicht an der gesättigten Wasser-Luft-Grenzfläche ist ca. 1.5 nm für alle beobachteten Tenside (siehe Abbildung 4.9). Somit variiert die Dichte der lyophilen Fragmente von der niedrigsten Dichte für $C_{16}E_6$ und der höchsten Dichte für C_6E_6. Dieses Resultat stimmt mit dem oben beschriebenen Ergebnis überein, indem eine längere C_x-Alkylkettenlänge zu einer geringeren Konzentration der Tensidmoleküle an der gesättigten Wasser-Luft-Grenzfläche führt. Allerdings widerspricht dieses Verhalten den von Rosen beschriebenen Ergebnissen[174]: Bei einer vollständig gesättigten Oberfläche sind die experimentellen Zahlen der Tensidmoleküle pro nm^2 1.52 (C_6E_6), 1.82 ($C_{10}E_6$), 1.92 ($C_{12}E_6$) und 2.63 ($C_{16}E_6$). Dies impliziert den genauen gegenteiligen Trend, dass eine längere C_x-Alkylkettenlänge zu einer höheren Tensiddichte an der Grenzfläche führt. Genauer gesagt, resultiert die hydrophobe Interaktion zwischen den Alkylketten in einer parallelen Ausrich-

tung an der Grenzfläche, was nicht vollständig von der Simulation abgebildet wird. Vielmehr sind in der Simulation Schlaufen zu sehen, welche eine größere Fläche einnehmen als die geradlinige Ausrichtung.

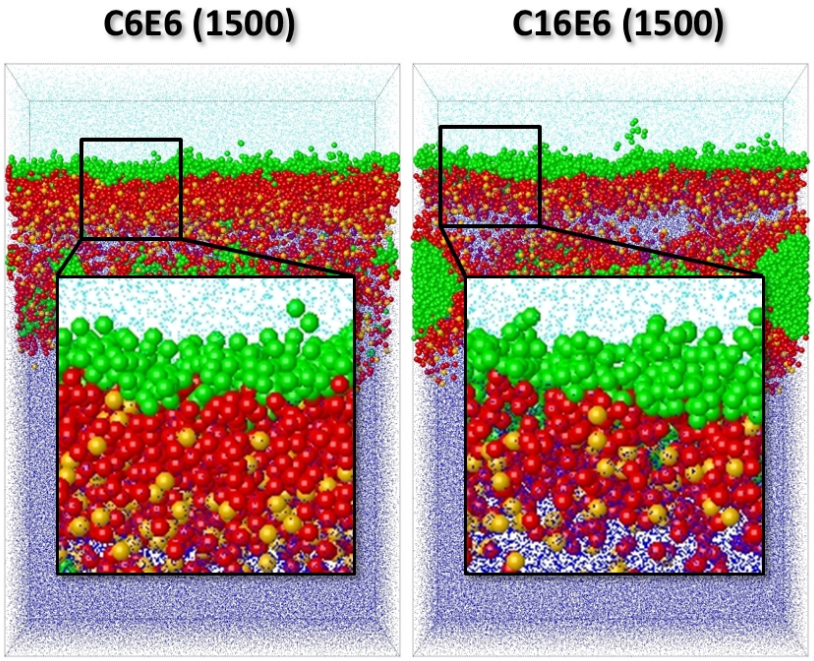


Abbildung 4.9: *Vergrößerte Detailansicht der gesättigten Wasser-Luft-Grenzfläche für C_6E_6 (links) und $C_{16}E_6$ (rechts)*[48].

5 Einfluss der effektiven Ladung auf die Aggregation von ionischen Systemen

In diesem Abschnitt wird der Einfluss der Stärke der effektiven Ladung q^* auf die Aggregation von Ionen betrachtet.

Methode

Die Untersuchung der effektiven Ladung q^* wird in einer quadratischen Simulationsbox mit einer Kantenlänge von 231.5 Å durchgeführt. Die Simulation enthält 10^5 Wassermoleküle und 1000 Methylaminacetatmoleküle (siehe Abbildung 5.1). Die periodischen Randbedingungen sind in alle Richtungen aktiviert. Tabelle 6.3 zeigt die gesetzten Parameter der Simulation. Untersucht wird der fragmentbasierte Gyrationsradius zur Beschreibung der Aggregation der untersuchten Teilchen.

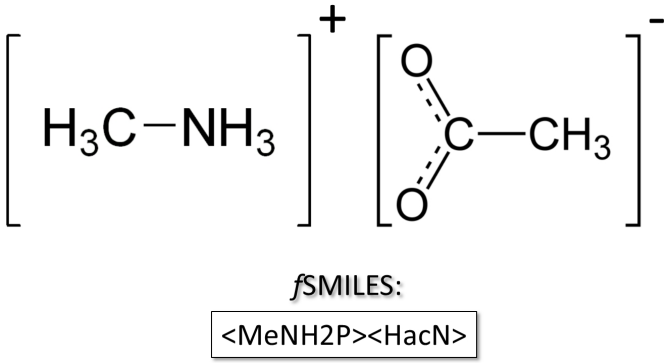


Abbildung 5.1: *Molekulare Struktur und fSMILES von Methylaminacetat.*

Ergebnis

Je eine Simulation des Korrekturfaktors der effektiven Ladung $c_{effective}$ mit den Werten 0.0, 0.1, 1.0, 2.0, 3.0, 4.0 und 5.0 werden durchgeführt. Die Ergebnisse sind in Abbildung 5.2 zusammengefasst, wobei die Simulationszeit gegen die Gyrationsradien des Fragmentpaares MeNH2P und HacN aufgetragen sind. Deutlich zu erkennen ist, dass die Entwicklung des Gyrationsradius für $c_{effective} \leq 1.0$ keine Veränderung gegenüber der Simulation mit ausgeschalteter elektrostatischer Kraft $c_{effective} = 0.0$ aufzeigt. Für $c_{effective} \geq 2.0$ ist deutlich zu sehen, dass der Gyrationsradius eine geringerer Steigung aufweist und bei höheren Werten dieses Verhalten zunimmt. Daraus folgt, dass die Erhöhung der effektiven Ladung q^* zu einer Aggregation der Ionen führt. Dieser Effekt ist für $c_{effective} = 5.0$ dermaßen stark, dass der Gyrationsradius nur 14.3 Å beträgt. Dies bedeutet, dass der überwiegende Anteil der Ionenpaare aggregiert ist und keine dissoziierten Ionen frei in der Simulationsbox diffundieren.

Tabelle 5.1: *Eingesetzte Simulationsparameter. In MFD-Einheiten, falls nicht anders angegeben .*

Parameter	Wert
Temperatur	310 K
Iterationen	5000
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5
k	2

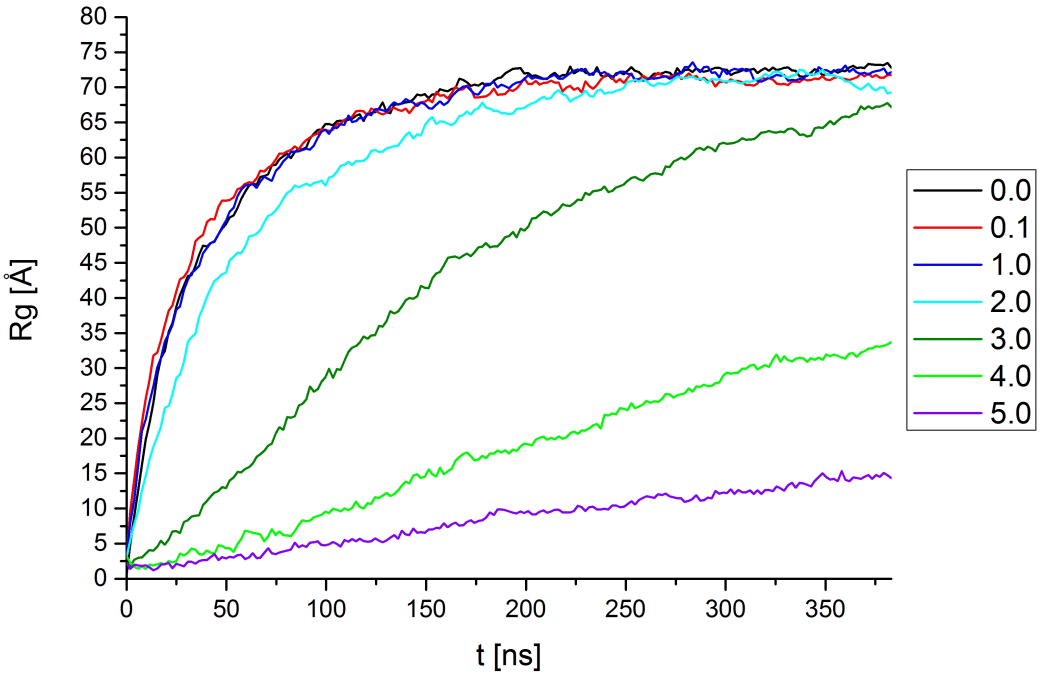


Abbildung 5.2: *Entwicklung der Gyrationsradien über die Zeit für verschiedene Korrekturfaktoren der effektiven Ladung* $c_{effective}$.

Die visuelle Begutachtung der Simulationsboxen im Gleichgewicht (siehe Abbildung 5.3) zeigt deutlich, dass schon für $c_{effective} = 1.0$ vereinzelt Ionenpaare entstehen. Für $c_{effective} \geq 2.0$ überwiegt die Anzahl der aggregierten Teilchenpaare und für $c_{effective} = 5.0$ ist jedes Kation mit je einem Anion gepaart. Freie Ionen sind an dieser Stelle nicht mehr erkennbar.

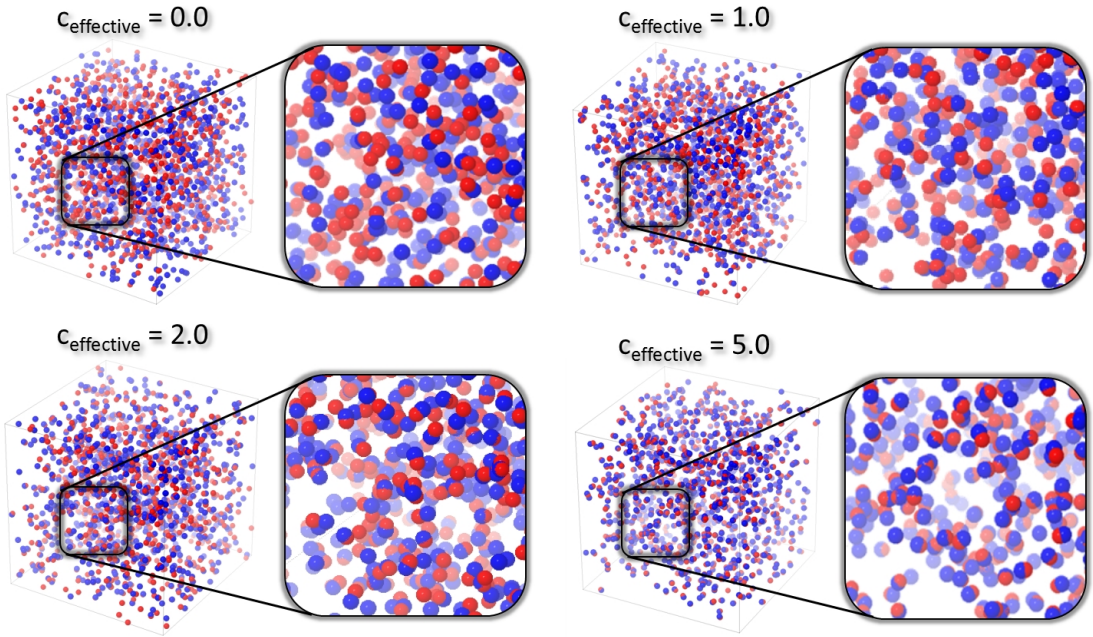


Abbildung 5.3: *Darstellung der Simulationsboxen im Gleichgewicht für verschiedene Korrekturfaktoren der effektiven Ladung* $c_{effective}$.

Die Clusterbildung der Ionen ist ein zu vermeidender unerwünschter künstlicher Effekt[67]. Aus den hier durchgeführten Simulationen geht hervor, dass ein Wert von $c_{effective} = 1.0$ zu einer realistische Verteilung zwischen aggregierten Teilchen und freien Teilchen führt. Der so ermittelte Wert wird für alle folgenden Versuche angewendet.

6 MFD-Studien von DMPC und biologischen Membranen

In diesem Kapitel wird die Einführung von Lipiddoppelschichten insbesondere von biologischen Membranen (Plasmamembran, Mitochondriummembran und die Membran des endoplasmatischen Retikulums) in die molekulare Fragmentdynamik betrachtet. Hierzu gehört die Untersuchung von Kenngrößen wie der Fläche pro Lipid, der hydrophoben Schichtdicke sowie der Gesamtschichtdicke, des Transmembranaustausches von Lipiden, der Mizellenformierung, der spontanen Bildung von Vesikeln und der Fusion einer DMPC-Membran mit einem Vesikel.

6.1 Biologische Membranen in molekularen Fragmentdynamik

Einfache Modellmembranen wie Dimyristoylphosphocholin-(DMPC)-Doppelschichten bestehen aus nur einer einzelnen Lipidspezies. Membranen des endoplasmatischen Retikulums (ER), von Mitochondrien sowie Plasmamembranen haben eine weitaus komplexere Struktur. Die häufigsten Lipide sind die Glycerophospholipide Phosphatidylcholin (PC), Phophatidylethanolamin (PE), Phosphatidylinositol (PI) und Phosphatidylserin (PS). Hinzukommen noch Sphingomyeline (SM) und Cardiolipin (CL). Die Zusammensetzung der Membranen der hier durchgeführten Studien sind in Abbildung 6.1 dargestellt und sind von van Meer et al.[76] abgeleitet. Die Membran des ERs setzt sich aus 50 % PC, 30 % PE, 15 % PI und 5 % PS zusammen. Die mitochondriale

Membran ist aus 50 % PC, 30 % PE, 10 % PI und 10 % CL aufgebaut. Die Plasmamembran besteht aus 40 % PC, 20 % PE, 5 % PI, 10 % PS und 25 % SM. Alle biologischen Membranen in dieser Studie sind wie hier beschrieben zusammengesetzt. Für die Lipidspezies Phosphatidylcholine (PC) werden vier verschiedene Varianten verwendet: 1,2-Dimyristoyl-sn-glycero-3-PC (DMPC), 1,2-Dipalmitoyl-sn-glycero-3-PC (DPPC), 1,2-Dioleoyl-sn-glycero-3-PC (DOPC) and 2-Oleoyl-1-palmitoyl-sn-glycero-PC (POPC). Sie unterscheiden sich ausschließlich in Länge und Struktur der Fettsäuren.

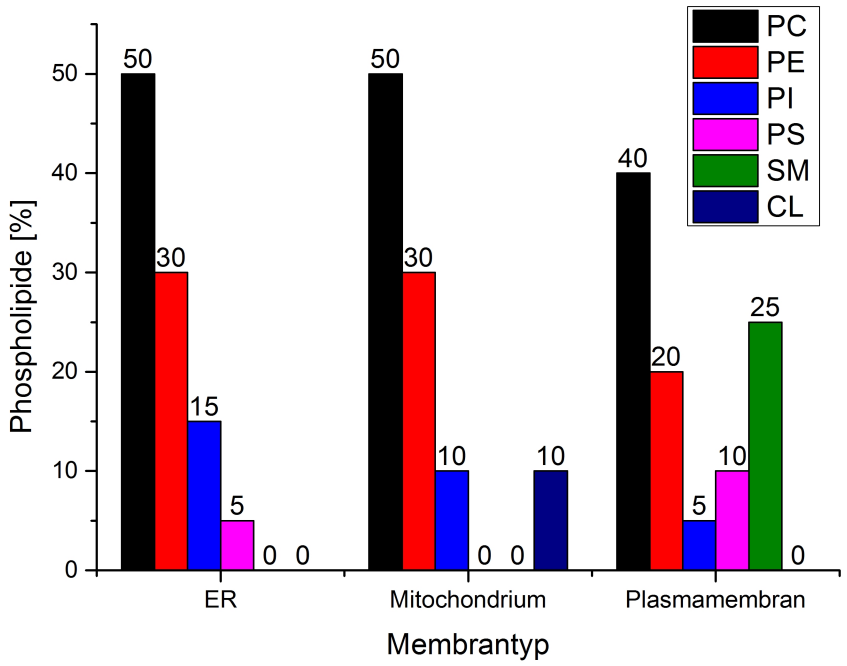


Abbildung 6.1: *Darstellung der Zusammensetzung der verschiedenen Membrantypen: ER, Mitochondrium und Plasmamembranen.*

Das Fragmentierungsschema aller Membranlipide ist in Tabelle 6.1 zu sehen. Abbildung 6.2 zeigt eine detaillierte Darstellung des Fragmentierungsschemas von Cardiolipin, welches nur in der mitochondrialen Membran vorkommt. Cardiolipin ist wegen der vier vorhandenen Fettsäurereste eine Ausnahme unter den Phospholipiden. Jedes Lipid besitzt eine Kopf- (*START*) und eine Schwanzmarkierung (*END*), um die Ausrichtung der Moleküle in der Doppelschicht der Startgeometrie festzulegen.

Tabelle 6.1: *Fragmentierungsschema der verwendeten Membranlipide.*

Membranlipid	Abkürzung	Fragmentierungsschema
1,2-Dimyristoyl-sn-glycero-3-PC	DMPC	TriMeNH2P[START]-DMPN(MeAc-6Et)(MeAc-6Et[END])
1,2-Dipalmitoyl-sn-glycero-3-PC	DPPC	TriMeNH2P[START]-DMPN(MeAc-8Et)(MeAc-8Et[END])
1,2-Dioleoyl-sn-glycero-3-PC	DOPC	TriMeNH2P[START]-DMPN(MeAc-Et-Pr-CisButene-3Et-Pr)(MeAc-Et-Pr-CisButene-3Et-Pr[END])
2-Oleoyl-1-palmitoyl-sn-glycero-PC	POPC	TriMeNH2P[START]-DMPN(MeAc-8Et)(MeAc-Et-Pr-CisButene-2Et-Pr[END])
2-Oleoyl-1-palmitoyl-sn-glycero-PE	POPE	MeNH2P[START]-DMPN(MeAc-Et-Pr-CisButene-2Et-Pr)(MeAc-8Et[END])
1,3-bis(1',2'-dilinoleoyl-sn-glycero-3'-phospho)-sn-glycerol	CL	DMPN(MeAc-Et-Pr-CisButene-CisButene-2Et)(MeAc-Et-Pr-CisButene-CisButene-2Et)-MeOH[START]-DMPN(MeAc-Et-Pr-CisButene-CisButene-2Et)(MeAc-Et-Pr-CisButene-CisButene-2Et[END])
Phosphatidylinositol-4,5-bisphosphate	PIP2	DMPN[1](MeAc-8Et)(4CisButene-Et[START])-MeOH-DMPN-DMPN[END]-MeOH-MeOH[1]
1,2-Diacyl-sn-glycero-3-phospho-L-serine	PS	HAcN[START](MeNH2P)-DMPN(MeAc-9Et)(MeAc-9Et[END])
Sphingomyelin	SM	TriMeNH2P[START]-DMPN(MeAcNH2-7Et)(MeOH-CisButene-6Et[END])

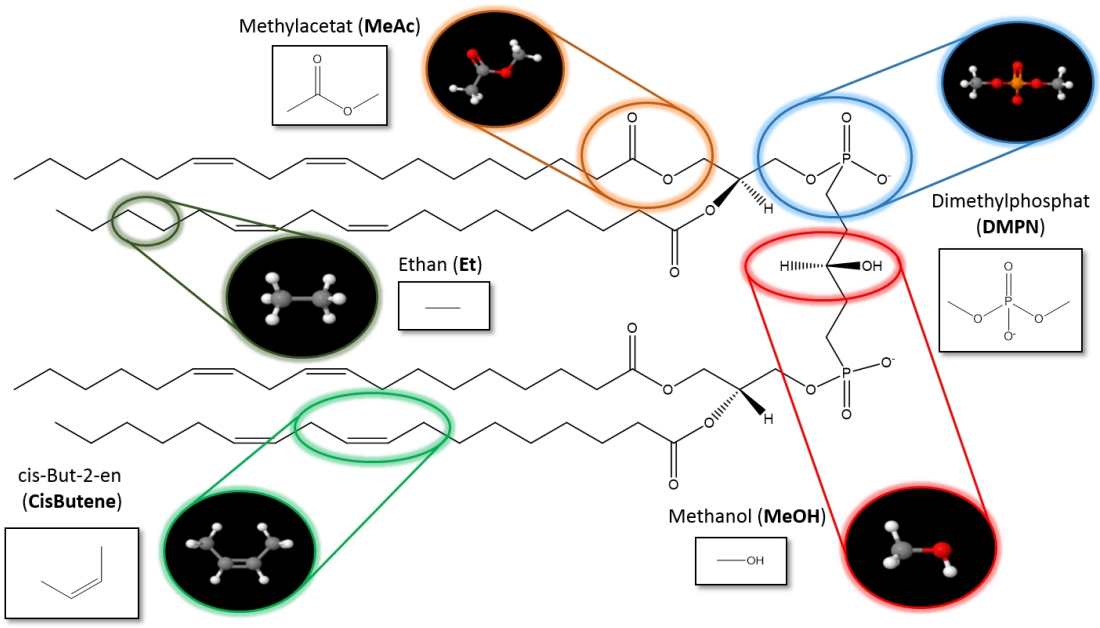


Abbildung 6.2: *Detailliertes Fragmentierungsschema von Cardiolipin. In den Klammern sind die jeweiligen Fragmentnamen abgebildet.*

Für eine leichtere Auswertung der Startgeometrien sowie der Ergebnisse zeigt Tabelle 6.2 die verschiedenen Färbungen der einzelnen Kopfgruppen der Membranlipide. Die hydrophoben Anteile sind stets grün gefärbt, soweit nicht anders in der Simulationsbeschreibung angegeben.

Tabelle 6.2: *Farbschema der Kopfgruppen der Membranlipide.*

Membrantyp	Farbe der Kopfgruppe
DMPC	blau (DMPN), rot (TriMeNHP)
DPPC	blau (DMPN), rot (TriMeNHP)
DOPC	blau (DMPN), rot (TriMeNHP)
POPC	blau (DMPN), rot (TriMeNHP)
POPE	gold (DMPN), orange (MeNH2P)
CL	cobalt (DMPN), indigo (MeOH)
PIP2	violett (DMPN), magenta (MeOH)
PS	oliv (HAcN, DMPN), mint (MeNH2P)
SM	braun (DMPN), grau (TriMeNHP)

Alle nachfolgenden Simulationen finden bei einer Temperatur von 310 K statt. Die periodischen Randbedingungen sind in alle drei Raumrichtungen aktiviert. Für die elektrostatische Kraft werden folgende Parameter eingesetzt:

Tabelle 6.3: *Simulationsparameter der elektrostatischen Kraft. In MFD-Einheiten, falls nicht anders angegeben.*

Parameter	Wert
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5

6.1.1 Einführung von DMPC und biologischen Membranmodellen

Methode

Für die Simulation von DMPC- und biologischen Membranen in der molekularen Fragmentdynamik werden kubische Simulationsboxen mit der Größe von 32·32·32 nm^3 konstruiert. Jede Box beinhaltet 2·10^5 Wassermoleküle. Für die Optimierung der DMPC-Membran werden Simulationen mit steigender Anzahl an Lipiden (1400 - 2400) wie in Tabelle 6.4 gezeigt gestartet. Die berechnete Fläche pro Lipid (Boxkantenlänge^2/Anzahl der Lipide) in Ångstrom reicht von 68.0 bis 43.0 Å^2. Die Lipidmoleküle werden in einer xy-Schicht mitten in der Simulationsbox arrangiert. Experimentelle Daten zeigen, dass die Fläche pro Lipid in

Tabelle 6.4: *Anzahl der Lipide mit der entsprechenden Fläche pro Lipid.*

Lipidanzahl	1400	1600	1700	1800	2000	2200	2400
Fläche pro Lipid [Å^2]	68.0	60.5	57.7	55.0	50.0	46.0	43.0

realen Membranen bei 60.7 ± 0.5 Å liegt[177]. Daraus folgt, dass für die Simulation der biologischen Membranen (ER, Mitochondrium, Plasmamembran) eine Anzahl von 1600 Lipidmolekülen eingesetzt werden muss, was wiederum einer ungefähren Fläche pro Lipid von 62 $Å^2$ für die biologischen Membranen entspricht. Die Simulationszeit beträgt 1.7 μs (20000 Iterationen). Abbildung 6.3 zeigt die Startgeometrien für alle vier untersuchten Membrantypen.

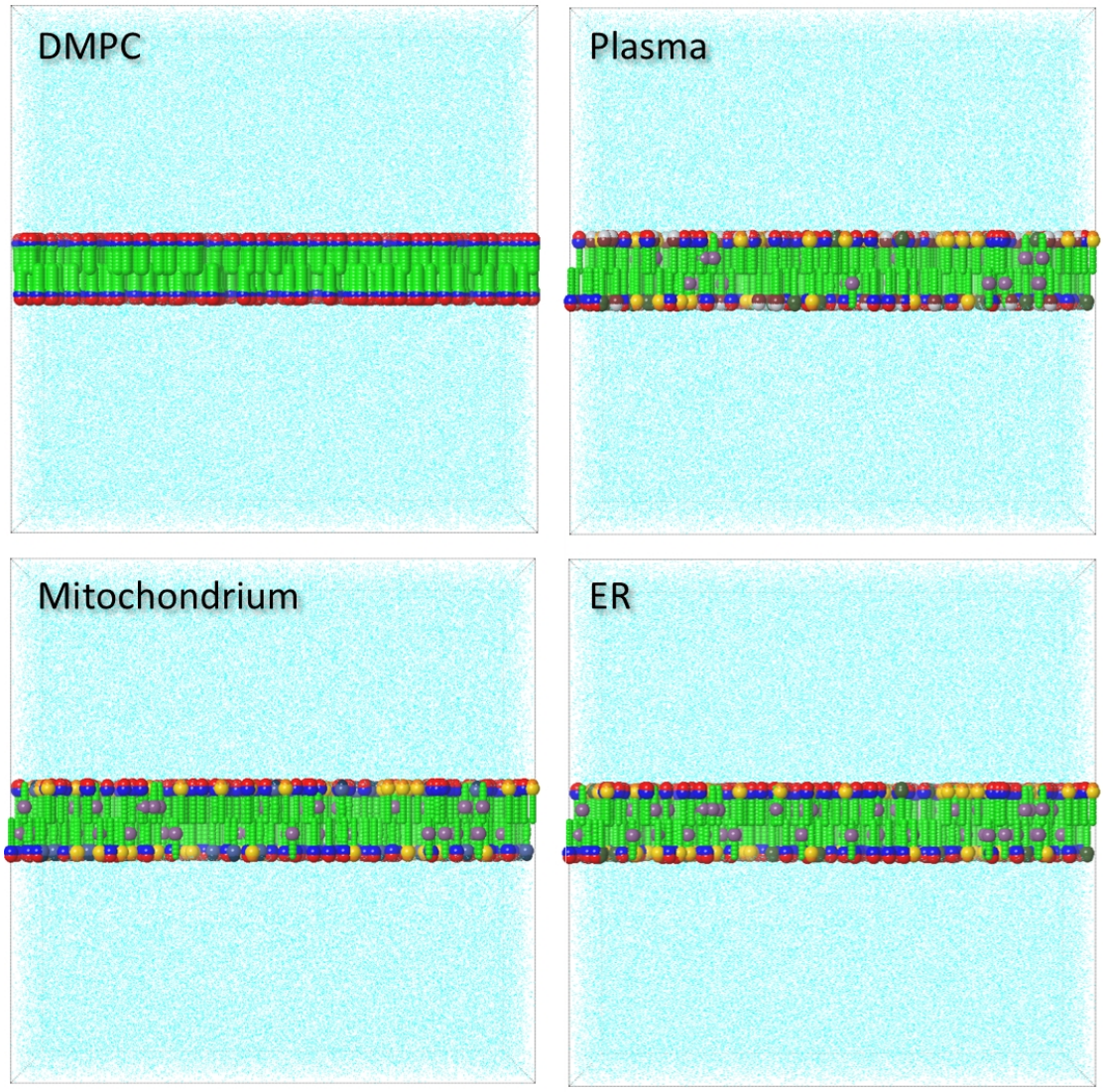


Abbildung 6.3: *Startgeometrie für alle vier untersuchten Membrantypen: DMPC, Plasma, Mitochondrium und ER.*

Ergebnis

Nach einer Simulationszeit von 1.7 μs werden die Dicke der hydrophoben Schicht sowie die Gesamtdicke der DMPC-Membran gemessen (siehe Abbildungen 6.4 und 6.5). Eine DMPC-Doppelschicht, bestehend aus 2400 Lipidmolekülen (Fläche pro Lipid: 43.0 Å^2), hat eine hydrophobe Dicke von 37.0 Å und eine Gesamtdicke von 58.0 Å. Ein Herabsetzen der Lipidanzahl auf 1400 (68.0 Å^2) resultiert in einer hydrophoben Dicke von 17.0 Å und einer Gesamtdicke von 33.0 Å. Die Gesamtbetrachtung der Daten zeigt, dass eine Anzahl an Lipiden von 1600 mit eine Fläche pro Lipid von 60.5 Å^2 die genaueste Übereinstimmung mit experimentellen Daten einer DMPC-Membran besitzt. Ein Querschnitt dieser DMPC-Membran ist in Abbildung 6.6 dargestellt.

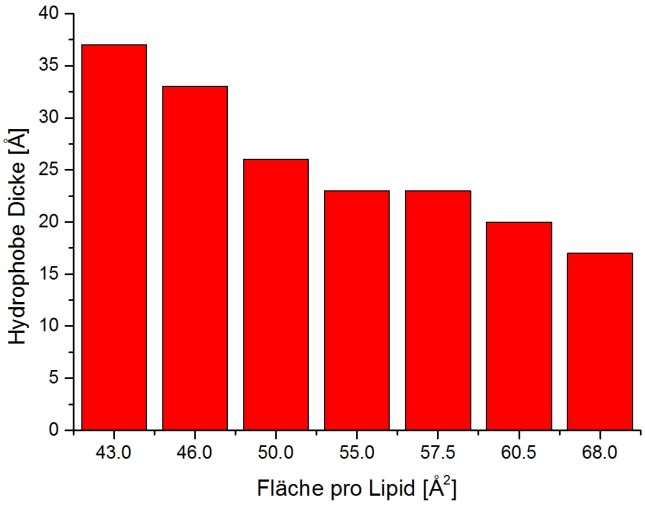


Abbildung 6.4: *Analyse der mittleren Dicke der hydrophoben Schicht für verschieden große Flächen pro Lipid.*

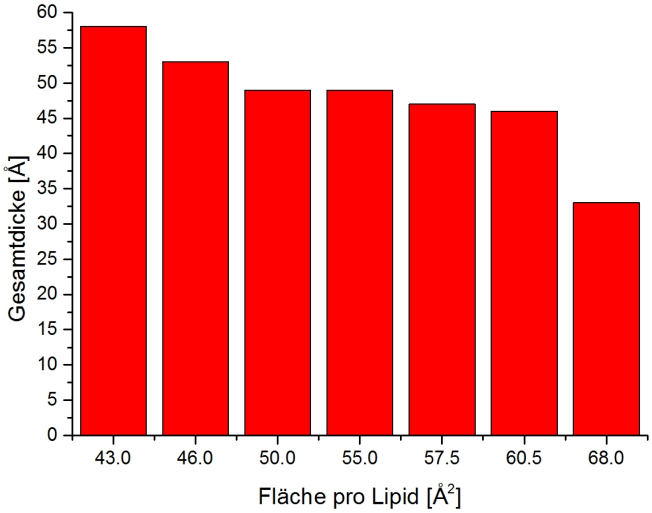


Abbildung 6.5: *Analyse der mittleren Dicke der gesamten Doppelschicht für verschieden große Flächen pro Lipid.*

Nicht nur die Fläche pro Lipidmolekül ist repräsentativ. Auch die hydrophobe Dicke (20 Å) und die Gesamtdicke (46 Å) stimmen mit den experimentell bestimmten Daten überein[177]. Das Erscheinungsbild der erhaltenen DMPC-Membran ist eine stabile Doppelschicht ohne Ausbildung von Löchern und Mizellen. Da für DMPC Membranen eine Lipidanzahl von 1600 zu einer angemessenen Beschreibung einer solchen Doppelschicht führt, wird dieselbe Anzahl an Lipiden für die Gestaltung der drei übrigen biologischen Membranen verwendet. Die Anteile der einzelnen Lipidmoleküle werden nach Tabelle 6.1 berechnet.

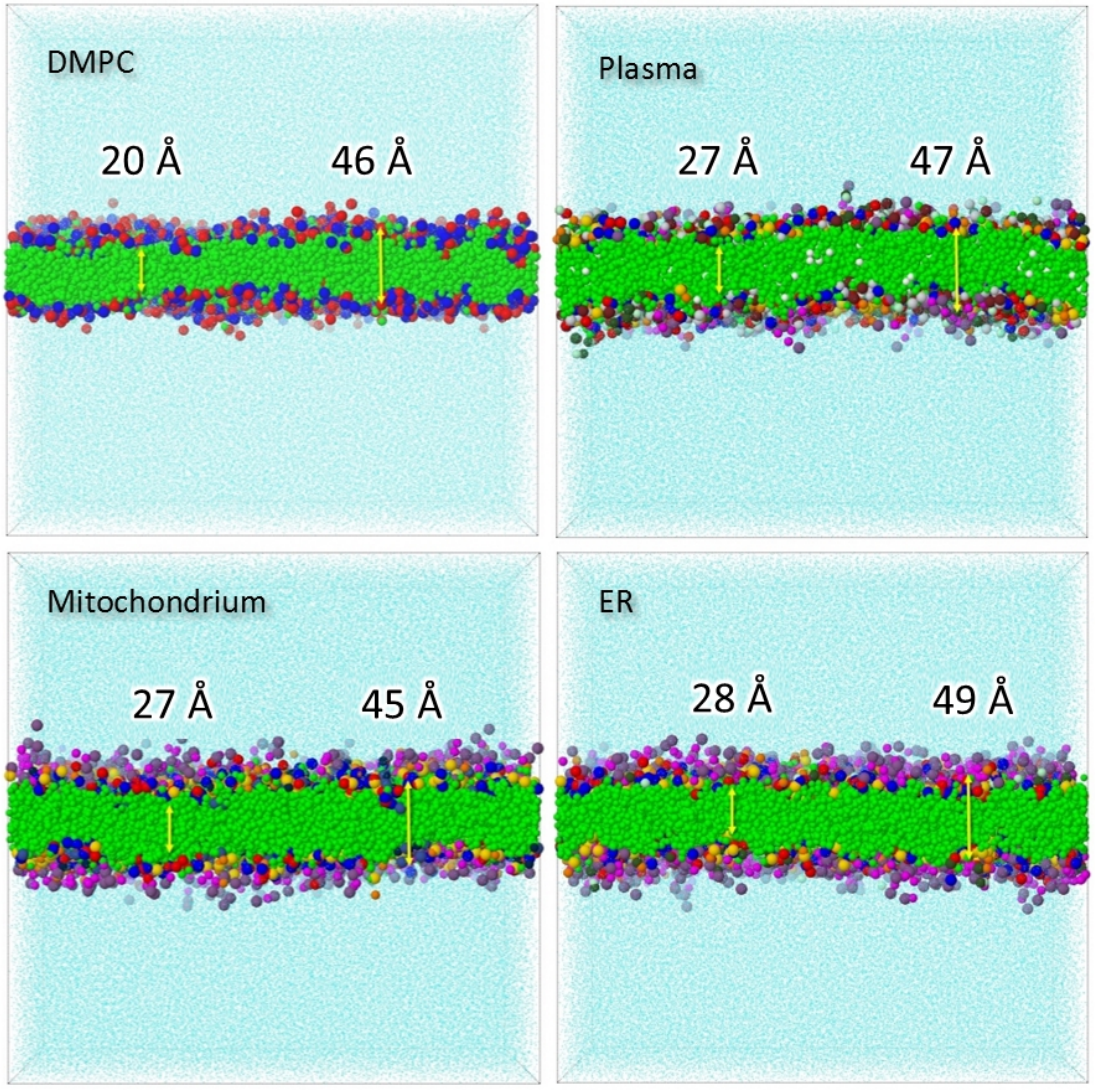


Abbildung 6.6: *Simulationsboxquerschnitte nach einer Simulationszeit von* 1.7 μs *aller vier Membrantypen.*

Zur Feststellung der hydrophoben Dicke und der Gesamtdicke wird die Membran an fünf verschiedenen Stellen stichprobenartig vermessen und der Mittelwert gebildet. Nach einer Simulationszeit von 1.7 μs werden folgende Ergebnisse der hydrophoben Dicke sowie der Gesamtdicke erhalten: Die Plasmamembran hat eine hydrophobe Dicke von 27.0 Å und eine Gesamtdicke von 47.0 Å. Die hydrophobe Dicke der Mitochondriummembran beträgt 27.0 Å und die Gesamtdicke 45.0 Å. Die Membran des endoplasmatischen Retikulums zeigt eine hydrophobe Dicke von 28.0 Å und eine Gesamtdicke von 49.0 Å. Alle drei Membranen bilden stabile Doppelschichten ohne Löcher und Mizellen aus. Alle Membranmodelle

zeigen wellenartige Fluktuationen an deren Oberflächen, wobei die drei biologischen Membranen insgesamt unruhiger wirken. Abbildung 6.6 zeigt Querschnitte der drei biologischen Membranen. Zusammengefasst weisen alle Membranmodelle eine gute Übereinstimmung mit experimentellen Daten auf.

6.1.2 Transmembranaustausch von Lipiden

Methode

Der Transmembranaustausch von Lipiden (auch als *Flip-flop* bezeichnet) ist eine weitere Eigenschaft von Membranen[139,140,178,179]. Die Betrachtung dieser Eigenschaft der simulierten Membranen wird in einer 32·32·32 nm^3 Simulationsbox mit $2{\cdot}10^5$ Wassermolekülen durchgeführt. Die DMPC und die biologischen Membranen werden aus 1600 Lipidmolekülen (Fläche pro Lipid: 62 Å) aufgebaut und in einer xy-Schicht zentral in die Simulationsbox platziert. Das in Kapitel 6.1 vorgestellte Farbschema der einzelnen Lipidmoleküle muss für diesen Versuch aufgegeben werden, um den Transmembranaustausch verfolgen zu können. Die Lipide der unteren bzw. oberen Schicht werden hierfür unterschiedlich gefärbt. Die Kopfgruppen der oberen Schicht sind blau und rot, die Fettsäuren grün gefärbt. Die unteren Lipide haben orange und gelbe Kopfgruppen sowie oliv gefärbte Fettsäuren. Auf diese Weise kann die Anzahl der Lipide in der jeweils anderen Schicht ausgewertet werden. Ein doppelter Austausch ist mit dieser Methode nicht detektierbar.

Damit *Flip-flops* nicht durch spontane Relaxation der Startgeometrie auftreten, werden 300 Energieminimierungsschritte vor der eigentlichen Simulation durchgeführt. Die Simulationszeit beträgt insgesamt 4.2 μs (50000 Iterationen)

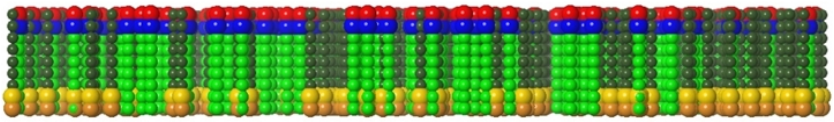


Abbildung 6.7: *Startgeometrie der Lipiddoppelschicht zur Untersuchung des Transmembranaustausches von Lipiden.*

Ergebnis

Nach einer Simulationszeit von 4.2 μs werden die Übergänge der Lipide innerhalb der Membranschichten ausgewertet. Mit einer Austauschrate von 37 Übergänge·μs^{-1} (156 Übergänge·4.2 μs^{-1}) zeigt die DMPC-Membran damit die höchste Anzahl an Übergängen pro Zeiteinheit. Die mitochondriale Membran folgt mit einer Übergangsrate von 28 Übergänge·μs^{-1} (123 Übergänge·4.2 μs^{-1}). Für die Membran des endoplasmatischen Retikulums wird eine Rate von 24 Übergänge·μs^{-1} (104 Übergänge·4.2 μs^{-1}) beobachtet. Die niedrigste Rate weist die Plasmamembran mit 17 Übergänge·μs^{-1} (73 Übergänge·4.2 μs^{-1}) auf. Im Vergleich mit DPD-Simulationen von DMPC-Membranen mit einer Austauschrate von 12 Übergänge·μs^{-1} [180] ist die Austauschrate der Lipide der hier betrachteten DMPC-Membran drei mal größer. Die experimentell bestimmte Verweildauer eines Phospholipids in seiner jeweiligen Schicht beträgt ungefähr 1 h [139,181]. Bezieht man diesen Wert auf das hier simulierte System, so kommt man auf eine Austauschrate von 4.4·10^{-7} Übergänge·μs^{-1}. Daraus folgt, dass die simulierten Membranen den Austausch von Lipiden reproduzieren, jedoch nicht die natürlich vorkommende Austauschrate. Abbildung 6.8 zeigt die Querschnitte der untersuchten Membranen.

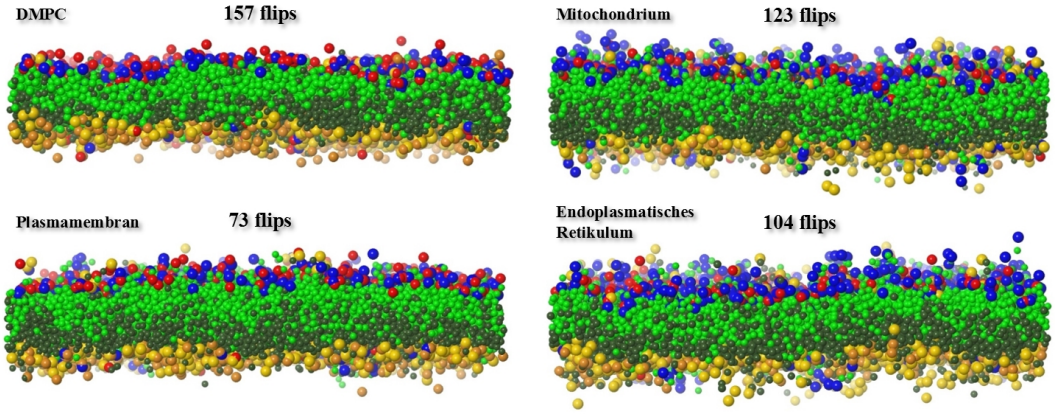


Abbildung 6.8: *Simulationsboxquerschnitte nach einer Simulationszeit von* 1.7 μs *aller vier Membrantypen.*

6.1.3 Mizellenformierung

Methode

Die Ausbildung von Mizellen ist eine Charakteristik von Phospholipiden[86,87]. Für die Untersuchung, ob die hier vorgestellten Modellmembranen fähig sind, Mizellen zu bilden, werden kubische Simulationsboxen mit einer Größe von 30·30·30 nm^3 mit $2{\cdot}10^5$ Wassermolekülen gefüllt. Die DMPC- und biologischen Membranen werden in einer zentralen xy-Schicht in der Simulationsbox platziert. Die Lipidanzahl wird von 1200 bis auf 400 in einer Schrittweite von 200 Lipide reduziert. Die Simulationszeit beträgt 1.5 μs (20000 Iterationen).

Tabelle 6.5: *Anzahl der Lipide mit der entsprechenden Fläche pro Lipid.*

Lipidanzahl	1200	1000	800	600	400
Fläche pro Lipid [$Å^2$]	78.0	92.0	112.0	147.0	216.0

Ergebnis

Die Ergebnisse für die DMPC-Membran sind in Abbildung 6.9 dargestellt. Die Simulationszeit beträgt 1.5 μs für alle abgebildeten Lipidkonzentrationen. Für eine Konzentration von 1200 bis zu 800 sind verzweigte Röhrenanordnungen zu erkennen, wobei mit abnehmender Lipidkonzentration Röhren mit geringerem Durchmesser entstehen. Bei einer Lipidanzahl von 600 finden sich drei Mizellen ohne Wassereinschluss wieder. Die Konzentration von 400 Lipidmolekülen führt zu einem Liposom mit Wassereinschluss.

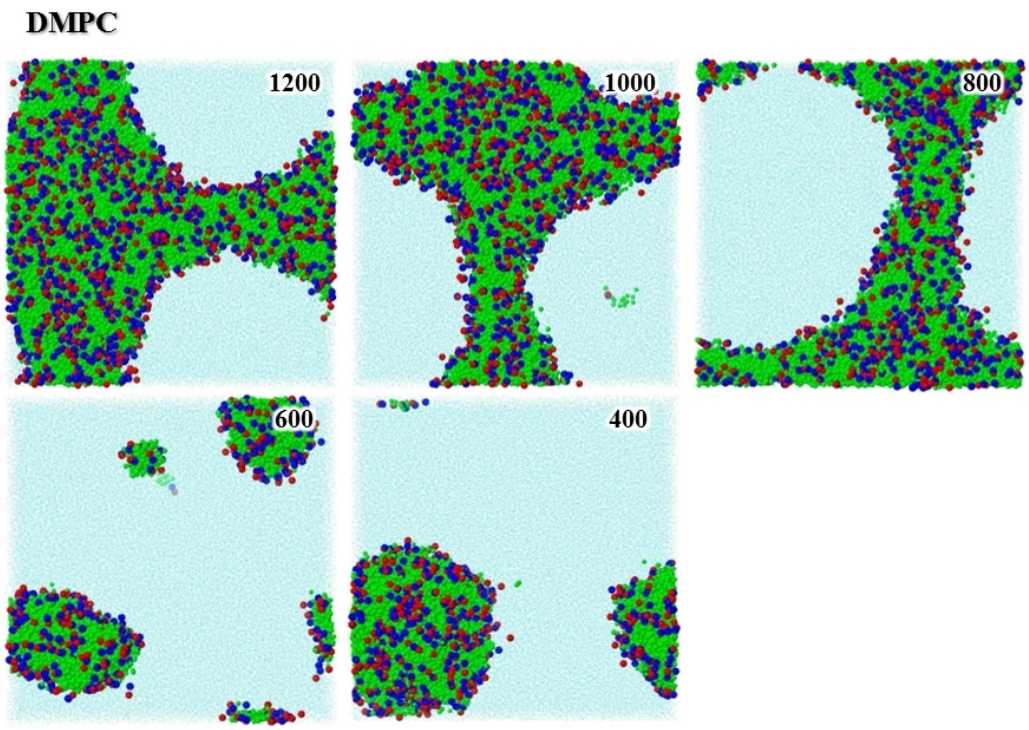


Abbildung 6.9: *Querschnitte der DMPC-Membran nach 1.5 µs, wenn die Lipidkonzentration von 1200 bis auf 400 Lipide gesenkt wird.*

Die MFD-Simulationen mit der Plasmamembran als Modelldoppelschicht erzeugen eine intakte Membran bei einer Lipidkonzentration von 1200 Molekülen. Keine Löcher oder Rupturen sind wahrnehmbar. Eine niedrigere Lipidkonzentration von 1000 bis 600 Molekülen resultiert in einem verzweigten Röhrensystem, wobei eine Konzentration von 600 Lipiden eine einzelne Röhre ausbildet. Bei der geringsten Lipidanzahl von 400 liegt ein Liposom mit minimalem Wassereinschluss vor. Abbildung 6.10 zeigt die Ergebnisse der Simulationen.

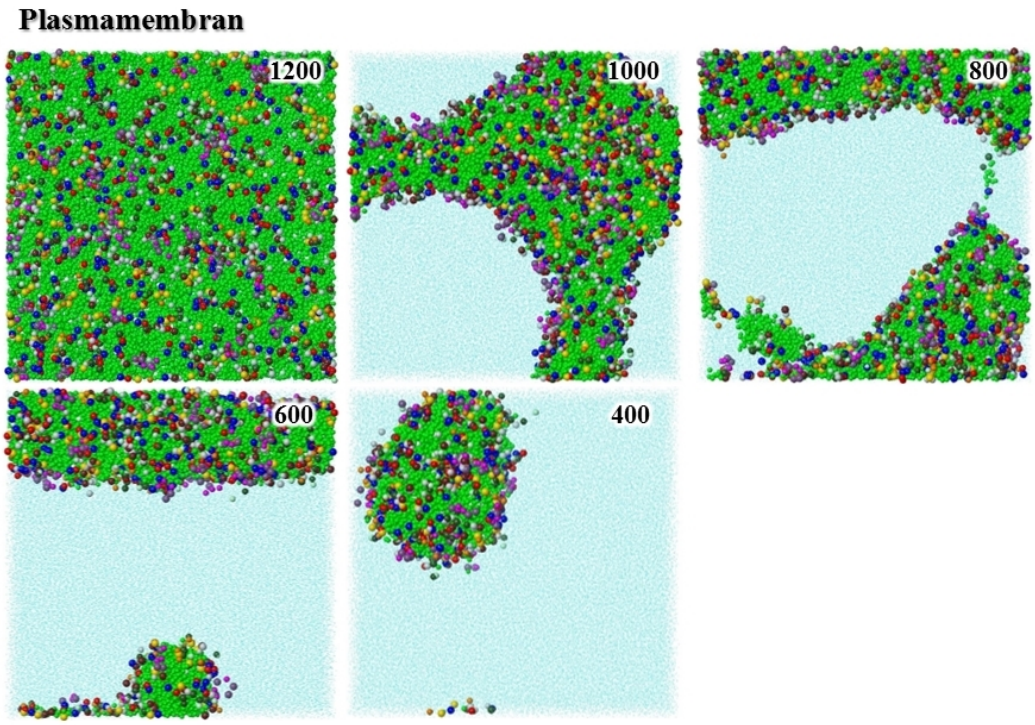


Abbildung 6.10: *Querschnitte der Plasmamembran nach 1.5 μs, wenn die Lipidkonzentration von 1200 bis auf 400 Lipide gesenkt wird.*

In Abbildung 6.11 sind die Ergebnisse der Mitochondriummembran nach einer Simulationszeit von 1.5 μs zu sehen. 1200 Lipidmolekülen bilden eine intakte Doppelschicht ohne Löcher oder Risse aus. Zwischen 1000 und 800 Lipidmolekülen reißt die Doppelschicht und ein großes rundes Loch entwickelt sich. Bei einer weiteren Verminderung der Lipidanzahl auf 600 zeigt sich eine einzelne Lipidröhre. Die niedrigste Konzentration (400) resultiert in zwei kleineren Mizellen und einem Liposom mit mehreren Wassereinschlüssen.

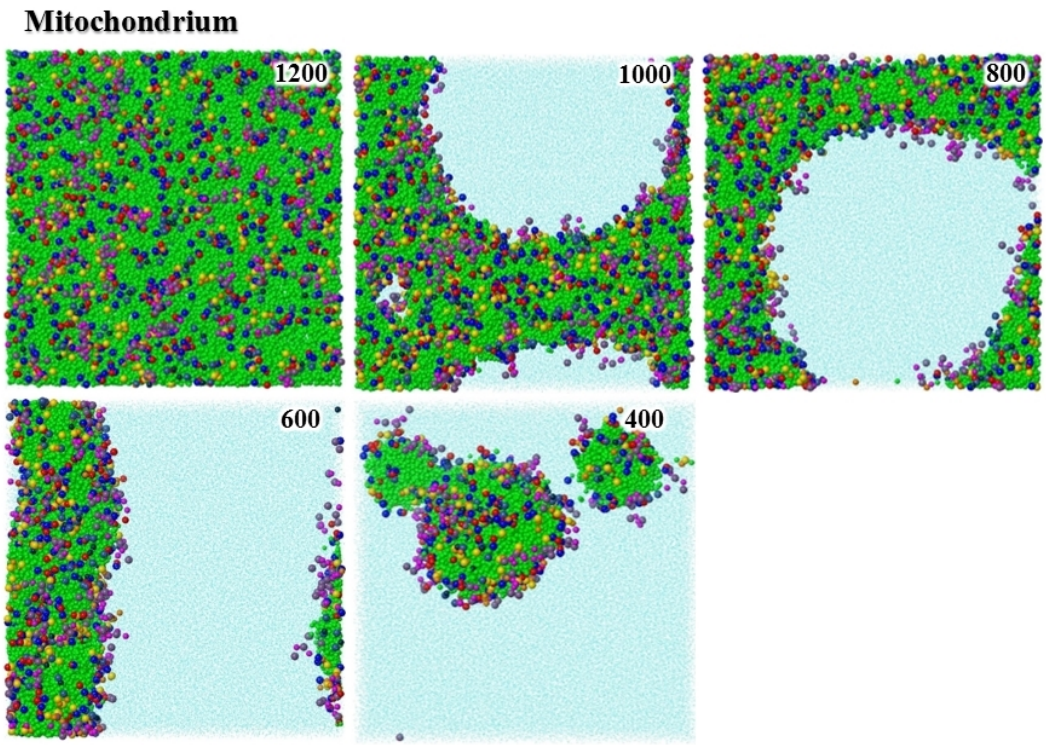


Abbildung 6.11: *Querschnitte der Mitochondriummembran nach 1.5 µs, wenn die Lipidkonzentration von 1200 bis auf 400 Lipide gesenkt wird.*

Die Untersuchung der Lipidkomposition für eine Membran des endoplasmatischen Retikulums ist in Abbildung 6.12 wiedergegeben. Die höchste Konzentration (1200) führt zu einer Doppelschicht mit rundem Loch. Das Herabsetzen der Konzentration auf 1000 bis 800 Moleküle führt zu verzweigten Lipidröhren mit geringer werdendem Durchmesser. Eine einzelne Lipidröhre entsteht bei einer Konzentration von 600 Lipiden. Die geringste Anzahl an Lipiden erzeugt eine Mizelle und ein Liposom mit geringem Wassereinschluss.

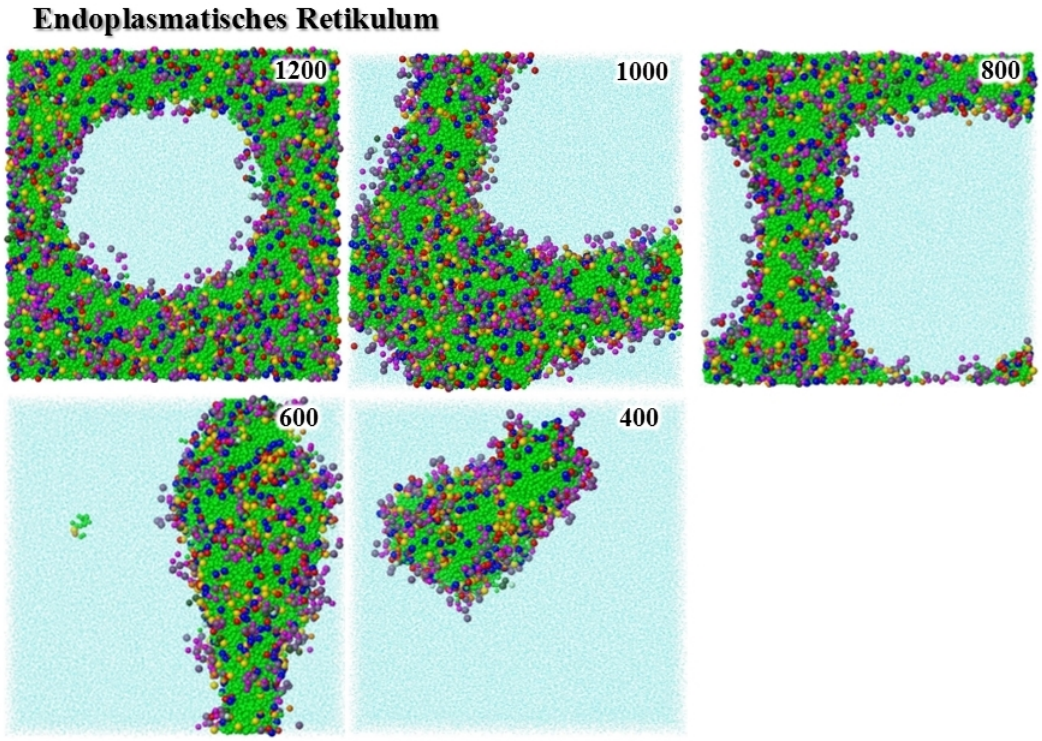


Abbildung 6.12: *Querschnitte der Membran des endoplasmatischen Retikulums nach 1.5 µs, wenn die Lipidkonzentration von 1200 bis auf 400 Lipide gesenkt wird.*

Eine Zusammenfassung der Ergebnisse zeigt: Die durchgeführten MFD-Simulationen erzeugen Mizellen und Liposomen wie in der Literatur beschrieben, falls die Konzentration von Lipidmolekülen nicht ausreicht, um stabile Membrandoppelschichten zu erzeugen. Die Ausbildung von verzweigten Lipidröhren wird in der Literatur nicht erwähnt.

6.1.4 Spontane Vesikelformierung

Methode

Eine weitere Studie betrachtet die Untersuchung der spontanen Vesikelbildung aller vier Membrantypen. Die Simulationsbox hierfür hat eine Größe von $35{\cdot}35{\cdot}35$ nm^3 und beinhaltet $3{\cdot}10^5$ Wassermoleküle. Die Anzahl der Lipidmoleküle beträgt für jeden Membrantyp 1800. Zwei unterschiedliche Startgeometrien werden

vorgegeben: (1) Eine zufällige Verteilung über die komplette Simulationsbox und (2) einen Doppelschichtblock in Form eines Ziegels im Zentrum der Box. Die Simulationszeit beträgt 16.40 μs ($2{\cdot}10^5$ Iterationen). Die Simulationsboxen wurden nach 0.08, 0.42, 2.40 und 16.40 μs analysiert.

Ergebnis

Phospholipide in wässriger Lösung sind bekannt dafür, dass sie Vesikel oder Liposomen ausbilden. Um die Selbstaggregation und Vesikelformierung untersuchen zu können, werden zwei unterschiedliche Startgeometrien angewendet: (1) Eine zufällige Anordnung der Lipide und (2) eine Anordnung in Form eines Ziegels ohne Kontakt zu den Wänden der Simulationsbox. Abbildung 6.13 zeigt die Ergebnisse der Vesikelformierung in einer zeitlich aufgelösten Bilderfolge für die zufällige Anordnung der DMPC-Moleküle. Nach 0.08 μs sind viele kleine sphärische Mizellen erkennbar. Bis 2.40 μs aggregieren die kleineren Mizellen zu immer größeren Strukturen. Die Bildung einer einzigen großen Mizelle ist im Zeitraum von 2.40 μs bis 16.4 μs verfolgbar. Nach Ausbildung der Mizelle fängt die Struktur zu fluktuieren an und verändert sich. Letztendlich entsteht ein DMPC-Doppelschichtvesikel mit Wassereinschluss (Abbildung 6.13, 16.40 μs - Querschnitt).

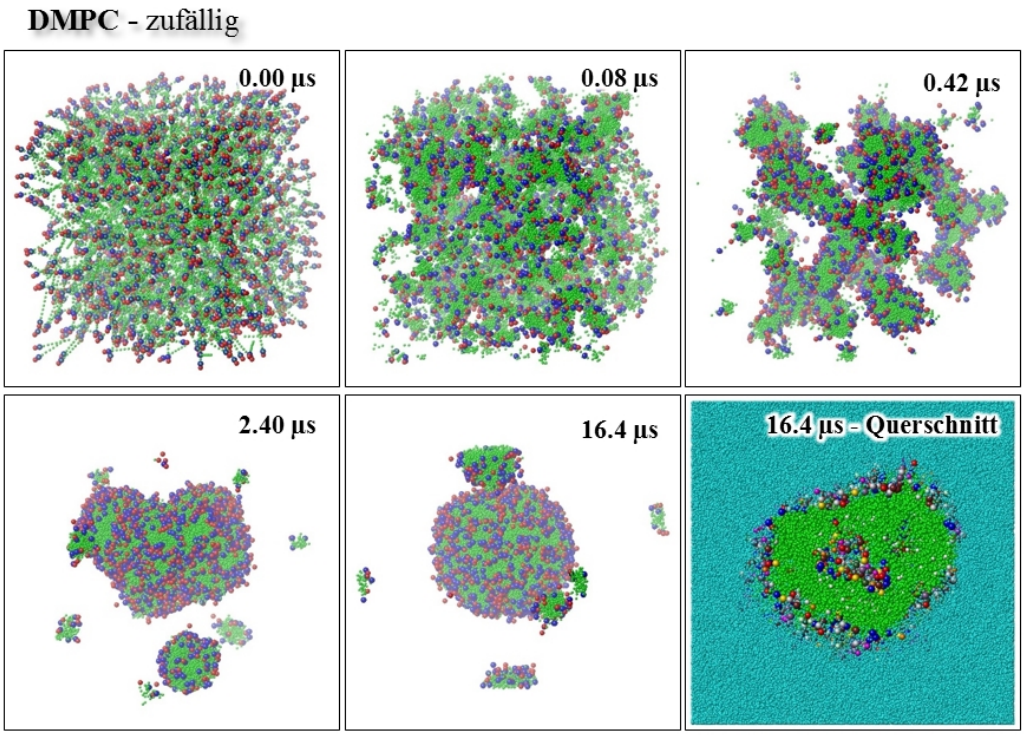


Abbildung 6.13: *Zeitliche Darstellung der Entwicklung einer zufälligen Anordnung von DMPC-Molekülen.*

Die Analyse der Ziegelstartgeometrie zeigt, dass die flache Ziegelstruktur bis 0.42 μs erhalten bleibt. Wassereinschlüsse sind bis zu diesem Zeitpunkt nicht aufzufinden. Ab 2.40 μs beginnt die Ausbildung einer mizellenartigen Struktur. Bis zum Ende der Simulation entsteht immer mehr eine DMPC-Doppelschichtstruktur. Zu sehen ist nicht nur ein Einschluss von Wassermolekülen wie bei der zufälligen Startgeometrie, sondern mehrere Zentren mit Wassereinschlüssen (Abbildung 6.14, 16.40 μs - Querschnitt).

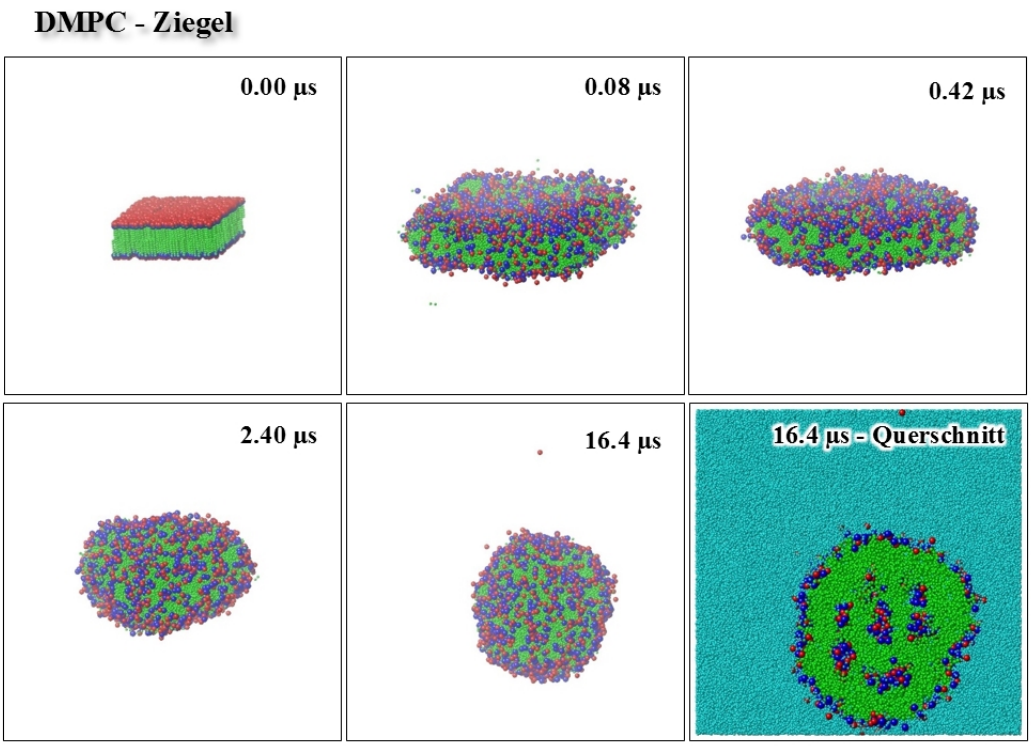


Abbildung 6.14: *Zeitliche Darstellung der Entwicklung der ziegelförmigen Startgeometrie der DMPC-Moleküle.*

Die gleiche Untersuchung wird mit den drei biologischen Membranen durchgeführt. Die Ergebnisse der Membran des Mitochondriums sind in den Abbildungen 6.15 und 6.16 dargestellt. Das Verhalten der Mitochondriummembran gleicht dem der vorher besprochenen DMPC-Membran. Am Anfang der Simulation bilden sich bei der zufälligen Startgeometrie viele kleine Mizellen aus, welche anschließend zu größeren Mizellen verschmelzen. Nach einer Simulationsdauer von 2.40 μs bildet sich eine einzige Mizelle aus, welche sich im weiteren Verlauf der Simulation in ein Wasser gefülltes Doppelschichtvesikel umwandelt.

Die flache ziegelartige Struktur bleibt bis 2.40 μs erhalten. Die immer runder werdende mizellenartige Struktur (2.40 μs) wandelt sich bis zum Zeitpunkt von 16.40 μs in ein Doppelschichtvesikel um. Im Gegensatz zu dem Ergebnis des DMPC-Versuches bildet sich hier nur ein Zentrum mit Wassereinschluss aus. Die beiden anderen Membrantypen zeigen dasselbe Verhalten wie die mitochondriale Membran.

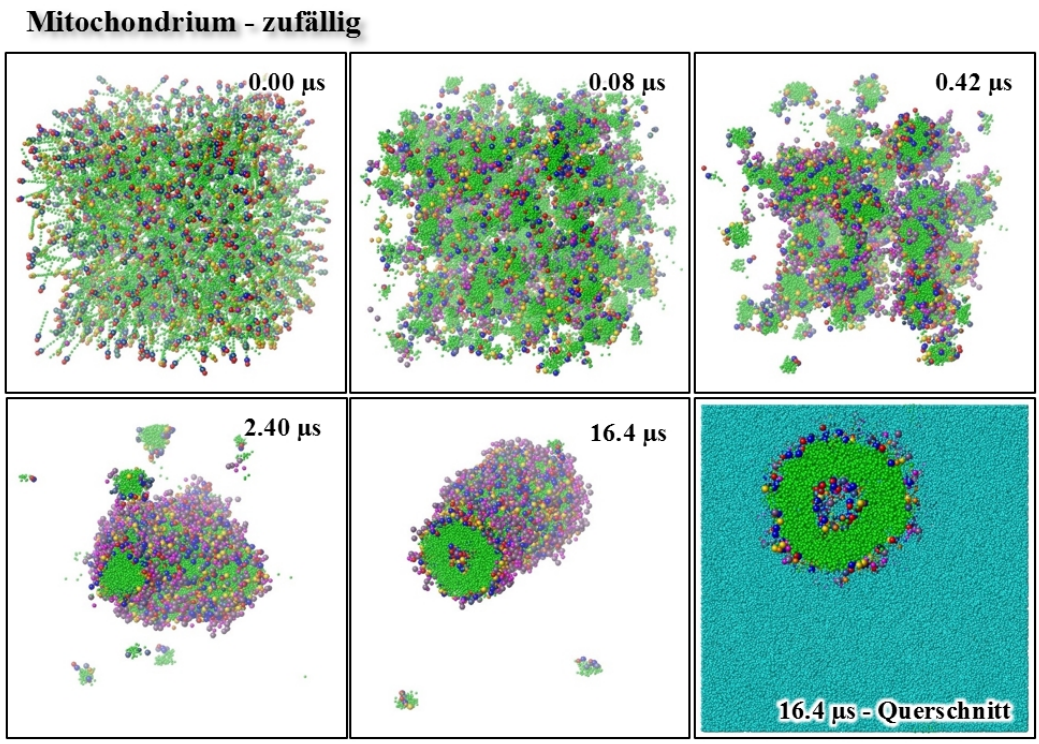


Abbildung 6.15: *Zeitliche Darstellung der Entwicklung einer zufälligen Anordnung der Lipide der mitochondrialen Membran.*

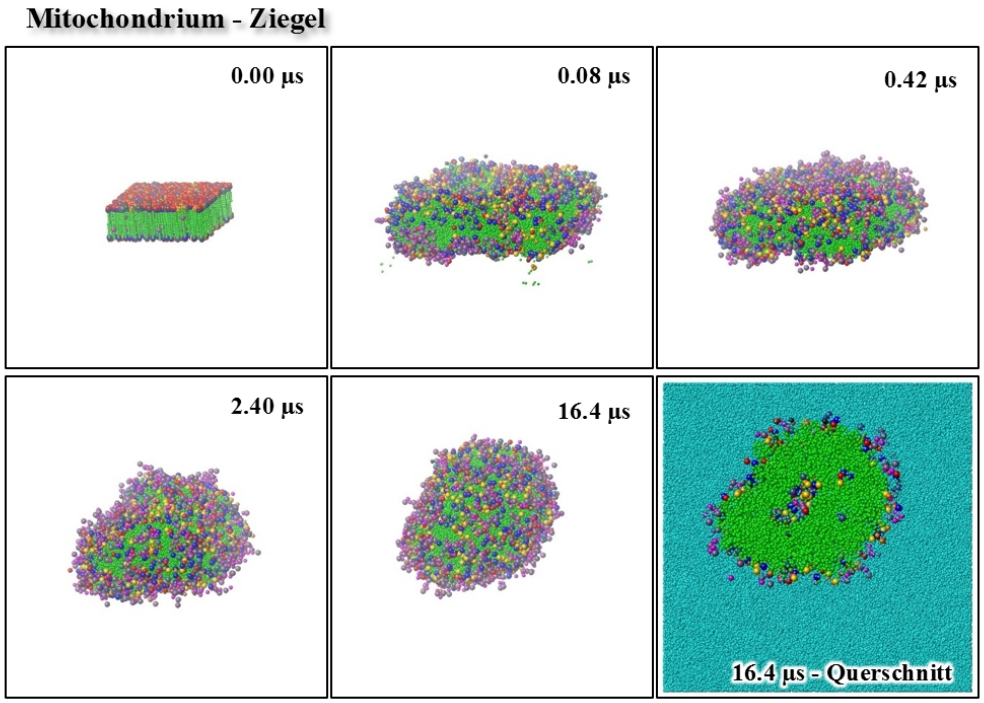


Abbildung 6.16: *Zeitliche Darstellung der Entwicklung der ziegelförmigen Startgeometrie der Lipide der mitochondrialen Membran.*

Zusammenfassend sei gesagt: Die Vesikelformierung, ausgehend von der zufälligen Anordnung der Lipide, dauert etwas länger als bei der ziegelartigen Startgeometrie. Die durchgeführten Versuche bestätigen, dass MFD-Simulationen Einsicht in die Mechanismen der Vesikelformierung von Phospholipiden geben.

6.1.5 Fusion einer DMPC-Membran mit einem Vesikel

Methode

Die Fusionierung von Phospholipidmembranen ist ein weiterer wichtiger Prozess für das biologische Leben. Hierzu gehören Prozesse wie der zellulare Transport mittels Vesikeln oder die Signalübertagung an den Synapsen.[88]. Um zu untersuchen, ob ein einfaches DMPC-Vesikel fähig ist mit einer DMPC-Doppelschichtmembran zu fusionieren und seine Fracht auf der anderen Seite der Membran zu entladen, wird eine große Simulationsbox mit $51{\cdot}51{\cdot}51$ nm^3 erstellt. Die Box besteht aus $8{\cdot}10^5$ Wassermolekülen sowie einer DMPC-Membran mit 4000 Lipidmolekülen. Die Membran ist im oberen Drittel positioniert. Die äußere Schicht des Vesikels ist aus 2200 und die innere Schicht aus 1800 Lipidmolekülen aufgebaut. Beide Schichten sind in zwei unterschiedliche Sphärenkompartimente untergebracht. Innerhalb der Vesikelmembransphären wird eine dritte Sphäre platziert. Diese Sphäre besteht aus $4.5{\cdot}10^4$ Wassermolekülen und bildet die Fracht des Vesikels. Der Durchmesser des Vesikels beträgt 30 nm und wird ohne Kontakt direkt unterhalb der Membran platziert. Das Vesikel wird unterschiedlich zur Membran eingefärbt, um die einzelnen Lipidmoleküle der Membran und des Vesikels voneinander unterscheiden zu können. Die Kopfgruppen der Membran sind blau, die Fettsäuren sind grün eingefärbt. Die Lipide des Vesikels tragen orange Kopfgruppen und die Fettsäuren sind gelb markiert. Das Wasser innerhalb des Vesikels ist rot koloriert. Die komplette Simulationszeit beträgt 3.4 μs ($4{\cdot}10^4$ Iterationen). Die Simulationsboxen wurden nach 0.01, 0.08, 0.47, 0.60, 0.94 und 3.40 μs analysiert.

Ergebnis

Für die Untersuchung des Prozesses der Vesikelfusion mit einer Membran wird eine stabile DMPC-Membran mit einem darunter liegenden mit Wasser gefülltem DMPC-Vesikel simuliert. Das Fusionsereignis wird ab dem Zeitpunkt des Kontaktes zwischen der Membran und des Vesikels beobachtet. Abbildung 6.17 zeigt mehrere Bilder des Fusionsprozesses aus einer Draufsicht und im Querschnitt. Nach 0.01 μs tritt das Vesikel mittels Diffusion in Kontakt mit der Membran. Durch den Zusammenstoß mit dem Vesikel wird die Membran nach oben gedrückt. Zu diesem Zeitpunkt ist noch kein Hinweis in der Draufsicht zu erkennen, dass das Vesikel im Kontakt mit der Membran steht. Nachdem sich die Kontaktfläche vergrößert (0.09 μs), tritt der erste Transmembranaustausch von Phospholipiden vom Vesikel in die Membran auf. Dies ist auch in der Draufsicht (gelbe und orange Punkte) zu sehen. Über Diffusionsprozesse gibt das Vesikel Wassermoleküle an seine Umgebung ab, was auch experimentell nachgewiesen ist[182]. Nach 0.47 μs wächst die Kontaktfläche weiter an. Erste Wassermoleküle diffundieren durch die DMPC-Doppelschicht. Transmembranübergänge von Lipiden können nun in beide Richtungen (Vesikel - Membran) beobachtet werden. Dies ist deutlich in der Draufsicht festzustellen. Zusätzlich ist ein nach außen gerichteter Knick an den Kontaktkanten zu sehen. Ab einer Simulationszeit von 0.60 μs ist deutlich die Entwicklung einer Pore wahrzunehmen. Eine erhöhte Menge an Wassermolekülen diffundiert durch die Pore auf die andere Seite der DMPC-Doppelschicht. Die Pore vergrößert sich bis 0.97 μs. Die hydrophoben Kerne des Vesikels und der Pore stehen im direkten Kontakt, aufgrund der Diffusion von Lipiden des Vesikels durch die hydrophobe Schicht in die obere Schicht der DMPC-Doppelschicht. Große Wassermengen fließen nun auf die andere Seite der Membran, welches zusätzlich in der Draufsicht zu erkennen ist. Am Ende der Simulation (3.40 μs) ist das Vesikel vollständig mit der Membran zu einer einzelnen Membran verschmolzen. Die Draufsicht zeigt, dass der überwiegende Anteil an Wassermolekülen, mit denen das Vesikel beladen ist, auf die andere Seite der DMPC-Doppelschicht entlassen wird. Aus alledem folgt, dass MFD-Simulationen fähig sind, die Fusion eines Vesikels mit einer Membran exakt wiederzugeben.

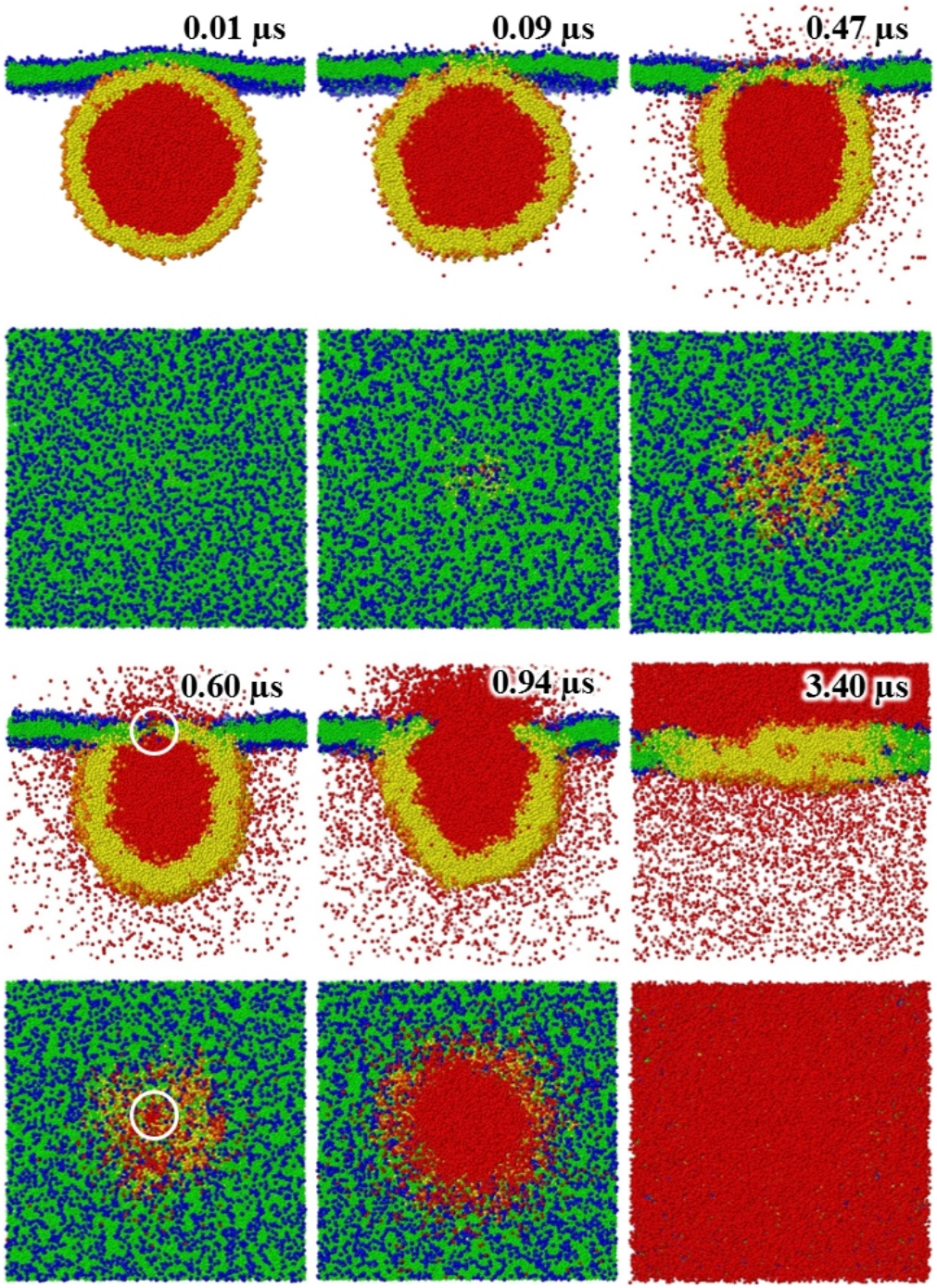


Abbildung 6.17: *Zeitlicher Verlauf der Vesikelfusion mit einer DMPC-Membran in einer Draufsicht sowie im Querschnitt.*

6.2 Diskussion

Viele computergestützte Simulationsmethoden wurden erfolgreich auf einfache Membranmodelle angewendet, um verschiedene Eigenschaften wie Stabilität, Zusammensetzung, Struktur und Fusion von Membranen zu studieren[79–82]. In dieser Arbeit konnte die MFD-Methode für Studien an einfachen DMPC-Membranen und erstmals auch an komplexen biologischen Membranen erfolgreich erprobt werden.

Die Modelle von DMPC-Membranen mit 1600 Lipiden in dieser Arbeit besitzen eine hydrophobe Dicke von 20.0 Å, eine Gesamtdicke von 46.0 Å und eine Fläche pro Lipid von 60.5 $Å^2$. Experimentelle Daten, die mittels Kleinwinkel-Neutronenstreuungsmethoden ermittelt wurden, zeigen, dass die hydrophobe Dicke 19.9 ± 0.4 Å und die Gesamtdicke 48.9 ± 0.2 Å für DMPC-Membranen beträgt. Die Fläche pro Lipid wird mit 60.7 ± 0.5 Å angegeben[183]. Des Weiteren ist das MFD-Membranmodell mit DPD-Studien von Grafmüller et al.[88] und Gao et al.[180] vergleichbar. Nicht nur die Fläche pro Lipid stimmt mit den experimentellen Daten überein, sondern auch die hydrophobe Dicke sowie die Gesamtdicke des hier entwickelten Modells.

Neben der Entwicklung eines Modells für einfache DMPC-Membranen konnten weitaus komplexere Modelle von biologischen Membranen wie die des endoplasmatischen Retikulums (ER), von Mitochondrien und Plasmamembranen erzeugt werden. Mit einer Fläche von 62.0 $Å^2$ pro Lipid ergibt sich eine hydrophobe Dicke von 27.0 Å und eine Gesamtdicke von 47.0 Å für die Plasmamembran. Die hydrophobe Dicke der Membran des Mitochondriums liegt bei 27.0 Å und die Gesamtdicke bei 47.0 Å. Die simulierte Membran des ER besitzt eine hydrophobe Dicke von 28.0 Å sowie eine Gesamtdicke von 49 Å. Verglichen mit experimentellen Daten, welche eine hydrophobe Dicke von 30 Å[76,83] und eine Gesamtdicke von 40 - 50 Å[84,85] aufweisen, liegen die drei simulierten biologischen Membranmodelle im gleichen Rahmen. Alle bilden stabile Doppelschichten ohne Löcher und Mizellenformierung. Die Oberflächen der Membranen zeigen wellenartige Fluktuationen mit gelegentlich auftretender Membrankrümmung wie auch in der Literatur

beschrieben[184].

Mit einer Rate von 37 Übergänge·μs^{-1} zeigt die DMPC Membran die höchste Transmembranaustauschrate von Lipiden. Es folgen die mitochondriale Membran mit 28 Übergänge·μs^{-1}, die Membran des endoplasmatischen Retikulums mit 24 Übergänge·μs^{-1} und die Plasmamembran mit der niedrigsten Rate von 17 Übergänge·μs^{-1}. Für die biologischen Membranen fehlen experimentelle Werte aufgrund von Limitierungen in der Bestimmungsmethode[140]. DMPC-Membranen sind jedoch besser charakterisiert, obwohl auch hier Unstimmigkeiten bzgl. des exakten Mechanismus sowie der Rate des Lipidaustausches vorliegen[179]. Im Vergleich mit DPD-Simulationen von DMPC-Membranen mit einer Austauschrate von 12 Übergänge·μs^{-1} liegt die MFD-Simulation drei mal höher. Die experimentell bestimmte Verweildauer eines Lipids in dessen aktueller Schicht reicht von 1 h[139,181] bis zu 1-2 min[179]. Demzufolge reproduzieren MFD-Membranen den Transmembranaustausch von Lipiden, aber nicht die natürlichen Übergangsraten. Eine Erklärung hierfür ist, dass der Transmembranaustausch von verschiedenen Membranproteinen kontrolliert bzw. beeinflusst wird, welche unter Energieaufnahme eine asymmetrische Verteilung der Lipide aufrecht erhalten[179,185–187]. Dieser Mechanismus sowie die zugehörigen Proteine sind noch nicht gänzlich charakterisiert, sodass MFD-Simulationen nur eine Energieminimierung der Modellmembranen durchführen.

Mizellen bestehen aus einer konzentrischen Lipiddoppelschicht ohne Einschluss. Liposomen hingegen sind mit einer wässrigen Phase gefüllt und umfassen eine oder mehrere konzentrische Doppelschichten. Die spontane Bildung dieser Strukturen ist eine allgemeine Eigenschaft von Phospholipiden[86,87]. Dieses Verhalten konnte auch für die vier Membrantypen in den hier durchgeführten Untersuchungen bei herabgesetzter Lipidanzahl beobachtet werden. Wenn die Lipidanzahl zu gering ist, um stabile Membranen zu bilden, so entstehen verzweigte und einfach röhrenförmige Membranen. Eine weitere Absenkung der Lipidanzahl führt zu der allgemein bekannten Bildung von Mizellen und Liposomen. Die Bildung von röhrenförmigen Strukturen wird nicht für Phospholipide in der Literatur beschrieben. In späteren Studien müsste untersucht werden, ob dieses Verhalten ein Artefakt der Simulati-

onsmethode ist.

Alle Membranen zeigen die Selbstorganisation von Vesikeln mit Wassereinschluss. Die Formierung der Strukturen dauerte für die zufällige Startkonfiguration länger, als jene für die Ziegelform. Die Simulationen der zufälligen Startgeometrie bilden anfangs viele kleinere Mizellen aus, welche zu einer großen Mizelle verschmelzen. Diese Mizelle unterliegt bis zum Ende der Simulationen konformellen Änderungen bis letztendlich Vesikel mit Wassereinschluss vorliegen. Die Ziegelstartgeometrie geht während der Simulation immer mehr in eine runde mizellenförmige Struktur über bis hin zu einem Vesikel mit mehreren Wassereinschlüssen.

Die Verschmelzung von biologischen Membranen ist ein wichtiger Prozess in lebenden Zellen. Dazu gehören Prozesse wie der zellulare Transport mittels Vesikeln oder die Signalübertragung an den Synapsen. Während des Prozesses der Verschmelzung unterliegt die Membran starken topologischen Veränderungen. Hierzu zählt ein Bruch der Doppelschicht, um eine erfolgreiche Verschmelzung des Vesikels und der Membran zu gewährleisten. Die Größe einer Verschmelzungspore liegt bei ungefähr 10 nm. Die geschätzte Zeitdauer der Verschmelzung ist kleiner 100 μs[188,189]. Zur Zeit gibt es keine experimentellen Möglichkeiten, dieses Ereignis innerhalb der angegebenen Größen- und Zeitskalen zu verfolgen[79,80,190–192]. Im Rahmen dieser Arbeit wird die Verschmelzung einer DMPC-Membran mit einem wassergefüllten DMPC-Vesikel untersucht. Das Vesikel hat einen Durchmesser von 30 nm. Nach dem ersten Kontakt des Vesikels mit der Membran dauert es 0.6 μs, um eine Verschmelzungspore zu bilden und eine massive Diffusion von Wasser durch die Membran einzuleiten. Die Größe der Pore beträgt ungefähr 10 nm entsprechend den Literaturdaten. Die gesamte Fusionszeit dauerte 3.4 μs. Der hier betrachtete Prozess stimmt mit der DPD-Simulation von Grafmüller et al.[88] überein. Lediglich der Zeitpunkt der Porenbildung von 1.3 μs liegt höher als in dieser Arbeit. Es konnte somit der Prozess der Vesikelverschmelzung mit einer MFD-Simulation erfolgreich dargestellt werden. Keine weiteren Arbeiten konnten gefunden werden, welche die komplette Vesikelverschmelzung darstellen.

7 Einführung von Peptiden und Proteinen in MFD

In diesem Kapitel werden alle Ergebnisse zur Simulation von Peptiden und Proteinen dargestellt. Es wird die Stabilisierung des Proteinrückgrats anhand des Beispielproteins Calmodulin untersucht. Die Aufrechterhaltung der Quartärstruktur von Hämoglobin wird für nur auf einzelne Untereinheiten beschränkte Stabilisierungspotentiale betrachtet. Nach der Auswertung der Peptid- und Proteinsimulationen wird eine Studie zur Interaktion von Zinkricinoleat mit histidinhaltigen Peptiden und Proteinen dargestellt. Am Ende wird das membranzerstörende Potential der Cyclotide, insbesondere des Proteins Kalata B1, unter Betrachtung der Wechselwirkung mit einer Plasmamembran, detailliert untersucht.

7.1 Stabilisierung des Protein-Rückgrates mittels harmonischer Potentiale

Die Sekundärstruktur von Proteinen wird durch Wassertstoffbrückenbindungen stabilisiert. Wasserstoffbrückenbindungen sind anisotrope Wechselwirkungen zwischen polaren Bereichen der molekularen Proteinstruktur. Jedoch werden MFD-Wechselwirkungen mittels des DPD-Repulsionsparameters beschrieben, welcher die chemischen Interaktionen der einzelnen Partikel in einem isotrop wirkenden Parameter annähert. Daraus folgt, dass eine Methode nötig ist, um die Proteinkonformation aufrecht zu erhalten. Die hier entwickelte Methode legt harmonische Potentiale auf die Fragmente des Proteinrückgrates, die die Steifigkeit

der aus der Kristallstruktur abgeleiteten Peptid- bzw. Proteinstruktur bestimmen. Über die Kraftkonstante der harmonischen Potentiale k_{BB} kann die Stärke der Stabilisierung frei gewählt werden.

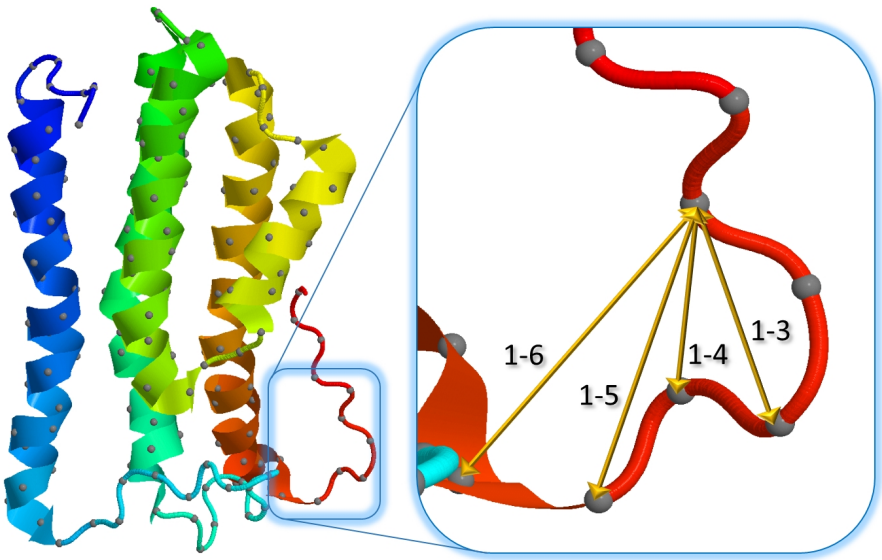


Abbildung 7.1: *Darstellung der auf die Fragmente des Proteinrückgrates (graue Kugeln) wirkenden harmonischen Potentiale (PDB-ID des dargestellten Proteins: 2LQU[193]).*

Die Anzahl der möglichen harmonischen Potentiale berechnet sich nach folgender Formel:

$$N_p = \frac{(N_{aa} - 1)N_{aa}}{2} \tag{7.1}$$

wobei N_{aa} die Anzahl der Aminosäuren im Peptid oder Protein ist. Die Nummerierung der Potentiale wird von der Entfernung der Fragmente in aufsteigender Richtung vom aminoterminalen Ende (N-Terminus) zum carboxylterminalen Ende (C-Terminus) zueinander abgeleitet. Das Fragment in Richtung des N-Terminus bekommt die Zahl 1, das zweite Fragment eine Nummer entsprechend dem Abstand zum ersten Fragment. Die Aminosäuresequenz des in Abbildung 7.1 dargestellten Proteins (PDB ID: 2LQU[193]) besteht aus 168 Aminosäuren. Aus Formel 7.1 ergibt sich eine Gesamtheit von 14028 Potentialen, die auf die Fragmente des Proteinrückgrates wirken und mit einer hohen Kraftkonstante die maxi-

male räumliche Stabilisierung ermöglichen. Im Gegensatz dazu folgt aus der Reduzierung der Anzahl der Potentiale und der Wahl einer kleinen Kraftkonstante eine flexible Proteinstruktur.

7.1.1 Auswertung der Stabilisierung des Protein-Rückgrates mittels harmonischer Potentiale

Als Studienobjekt soll das Protein Calmodulin (PDB-ID: 4BW8[194]) dienen. Calmodulin ist ein hochkonserviertes kleines Ca^{2+}-bindendes Protein mit einer Gesamtanzahl von 138 Aminosäuren. Es spielt eine große Rolle in vielen Ca^{2+}-abhängigen Signaltransduktionswegen. In Anwesenheit von Calcium hat Calmodulin die Form einer Hantel mit einem Gyrationsradius von $R_g = 21.5$ Å[195]. Die N- und C-terminalen Domänen sind durch eine 27-Aminosäuren lange α-Helix separiert. Jede der beiden homologen N- und C-terminalen Domänen beinhaltet zwei EF-Hand Motive. Jedes kann ein Ca^{2+}-Ion binden[194,196].

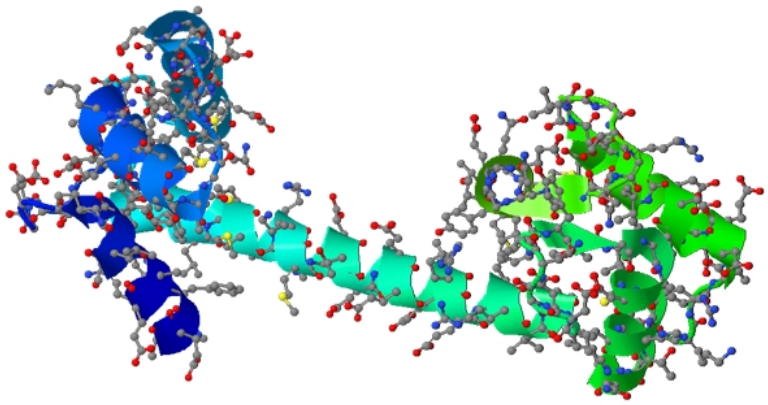


Abbildung 7.2: *Hantelförmige molekulare Struktur des Calmodulins.*

Methode

Die Simulationen finden in einer Box mit $5{\cdot}10^4$ Wasserfragmenten (H_2O) statt, in welcher das Calmodulin zentral positioniert wird. Die Ladungen des Proteins werden auf einen pH-Wert von 7.4 gesetzt. $2{\cdot}10^4$ Simulationsschritte pro Simu-

lation werden berechnet, was einer Zeitdauer von 1.5 μs entspricht. Die α-C-Fragmente 1, 50, 100 und 144 werden durch Sondenfragmente ersetzt. Die Box hat eine Größe von 182·182·182 Å^3. Abbildung 7.3 zeigt die Startgeometrie und Abbildung 7.4 die räumliche Anordnung der α-C-Sondenfragmente innerhalb des Calmodulins. Die Messungen der ausgewerteten Parameter starten nach 5000 Equilibrierungsschritten. Die Parameter der elektrostatischen Kraft sind in Tabelle 7.1 dargestellt.

Tabelle 7.1: *Simulationsparameter der elektrostatischen Kraft. In MFD-Einheiten, falls nicht anders angegeben.*

Parameter	Wert
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5

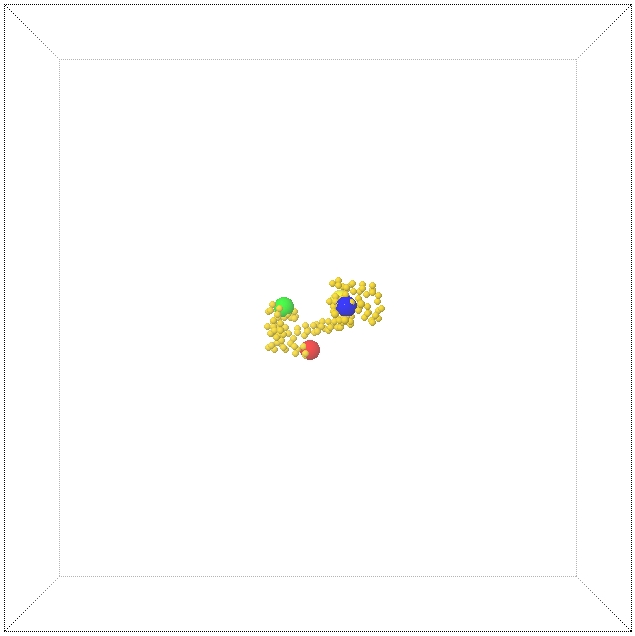


Abbildung 7.3: *Ansicht der Startgeometrie der Simulation. Zentral gelegen befindet sich das Calmodulin, wobei nur die α-C-Fragmente (gold) und die vier Sondenfragmente (rot, grün, blau, gelb) sichtbar sind. Für eine bessere Übersicht sind die Wassermoleküle ausgeblendet.*

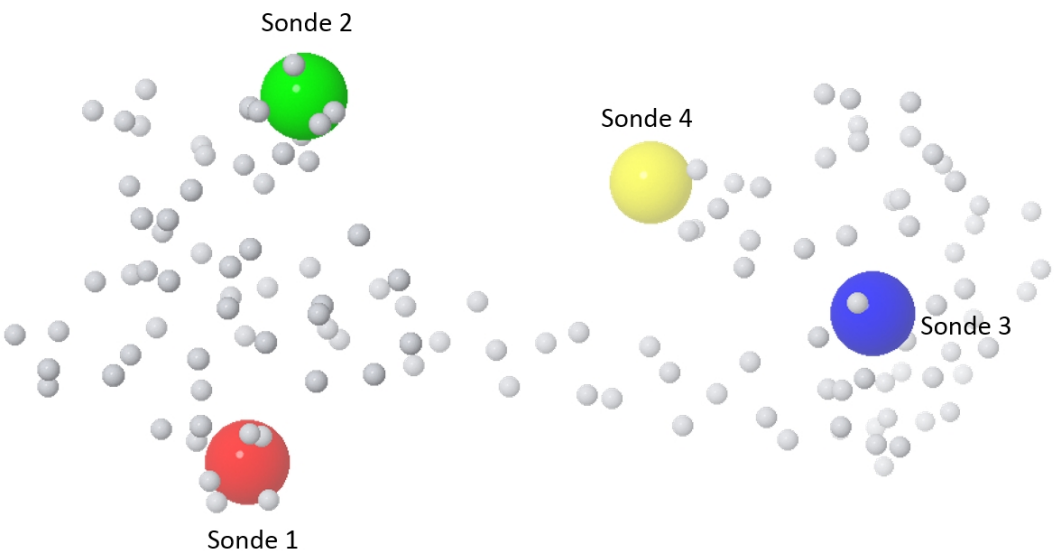


Abbildung 7.4: *Räumliche Anordnung der Sondenfragmente innerhalb des Calmodulin Proteinrückgrates. Grau dargestellt sind die restlichen α-C-Fragmente.*

In diesem Versuch werden vier Experimente durchgeführt. Die erste Simulationsreihe dient der Bestimmung des Einflusses der Kraftkonstante der stabilisierenden harmonischen Potentiale auf die Konformation des Proteins. Hierbei wird die Kraftkonstante k_{BB} von 0.0 bis 10.0 variiert. Es werden alle Potentiale über sämtliche Distanzen verwendet. Dies ergibt eine Gesamtheit von 10296 harmonischen Potentialen zur Stabilisierung des Proteinrückgrats. Mittels der übrigen drei Experimente soll der Einfluss der Länge der aktivierten Potentiale bei konstanter Kraftkonstante mit einem Wert von $k_{BB} = 2.0$ ermittelt werden. Untersucht werden die Auswirkungen der Potentiale mit kurzer Reichweite, langer Reichweite und einer Kombination aus Potentialen mit kurzer sowie langer Reichweite. Dabei sollen sukzessive kurze bzw. lange Potentiale hinzugefügt werden, um die Auswirkungen der einzelnen Potentiale zu charakterisieren. Tabelle 7.2 zeigt die Anzahl der Potentiale n_{BB} in Abhängigkeit der aktivierten Potentiale. Die Variable $d_{BB} = a - b$ beschreibt dabei die Länge der aktivierten Potentiale, d.h. bei einem Wert von 2-10 (ausgeschrieben: 1-2 bis 1-10) sind alle Potentiale mit einer Reichweite von ein bis neun α-C-Fragmenten aktiviert, d.h. alle Potentiale zwischen den α-C-Fragmenten (C^α) mit den Indices $C^\alpha_i - C^\alpha_{i+(a-1)}$ bis $C^\alpha_i - C^\alpha_{i+(b-1)}$ sind aktiv.

Tabelle 7.2: Anzahl der Stabilisierungspotentiale n_{BB} beim sukzessiven Hinzufügen von Potentialen mit:

(a) kurzer Reichweite.

d_{BB}	n_{BB}
0	0
2-11	1385
2-21	2670
2-32	3968
2-44	5246
2-57	6468
2-73	7740
2-94	9021
2-144	10296

(b) langer Reichweite.

d_{BB}	n_{BB}
0	0
144-94	1326
144-73	2628
144-57	3916
144-44	5151
144-32	6441
144-21	7750
144-11	9045
144-2	10296

(c) kurzer und langer Reichweite.

d_{BB}	n_{BB}
0	0
2-4, 144-109	1371
2-9, 144-94	2711
2-14, 144-83	3993
2-19, 144-73	5298
2-24, 144-65	6515
2-30, 144-57	7884
2-35, 144-50	9078
2-144	10296

Die vorgestellten Diagramme sind sogenannte Box-Whisker-Plots. Sie dienen der Darstellung von verschiedenen statistischen Kenngrößen der vorliegenden Daten (Blauer Kasten: Oberes und unteres Quantil; Weiße Linie: Der Mittelwert; Unterer und oberer *Whisker*: Kleinster bzw. größter Wert; Unterer bis oberer *Whisker*: Die Spannweite)[197]. Eine rote durchgezogene Linie zeigt die Lage des experimentell ermittelten Wertes, eine rote gestrichelte Linie entspricht dem Mittelwert aus der Simulation mit der höchsten Strukturkonservierung (höchste Kraftkonstante k_{BB} der harmonischen Potentiale).

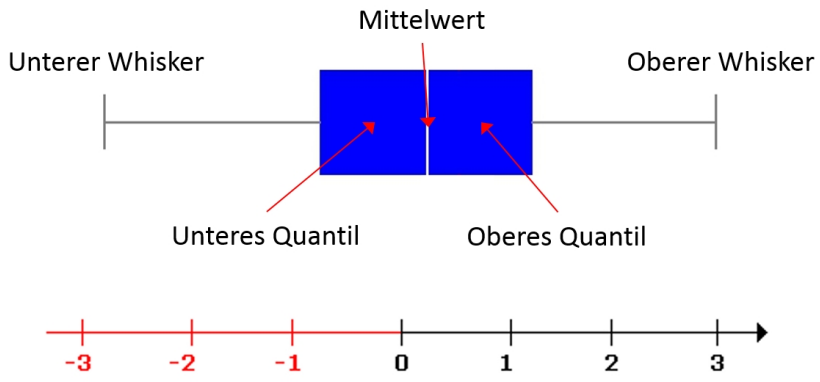


Abbildung 7.5: *Box-Whisker-Plot.*

Ergebnis

Die erste Simulationsreihe dient der Bestimmung des Einflusses der Kraftkonstante der stabilisierenden harmonischen Potentiale auf die Geometrie des Proteins. Hierbei wird die Kraftkonstante k_{BB} von 0.0 bis 10.0 variiert. Es werden alle Potentiale über alle Distanzen verwendet. Dies ergibt eine Gesamtheit von 10296 harmonischen Potentialen, welche das Proteinrückgrat stabilisieren.

Das Diagramm 7.6 zeigt den Einfluss der Stärke der Kraftkonstante k_{BB} auf den fragmentbasierten Gyrationsradius $R_{g,fragment}$. Es fällt auf, dass der Mittelwert des fragmentbasierten Gyrationsradius der Simulation größer ist als der eigentliche experimentell bestimmte Gyrationsradius. Der Grund dafür ist, dass MFD eine fragmentbasierte Berechnung des Gyrationsradius $R_{g,fragment}$ implementiert und keine atomistische Betrachtung zugrunde liegt.

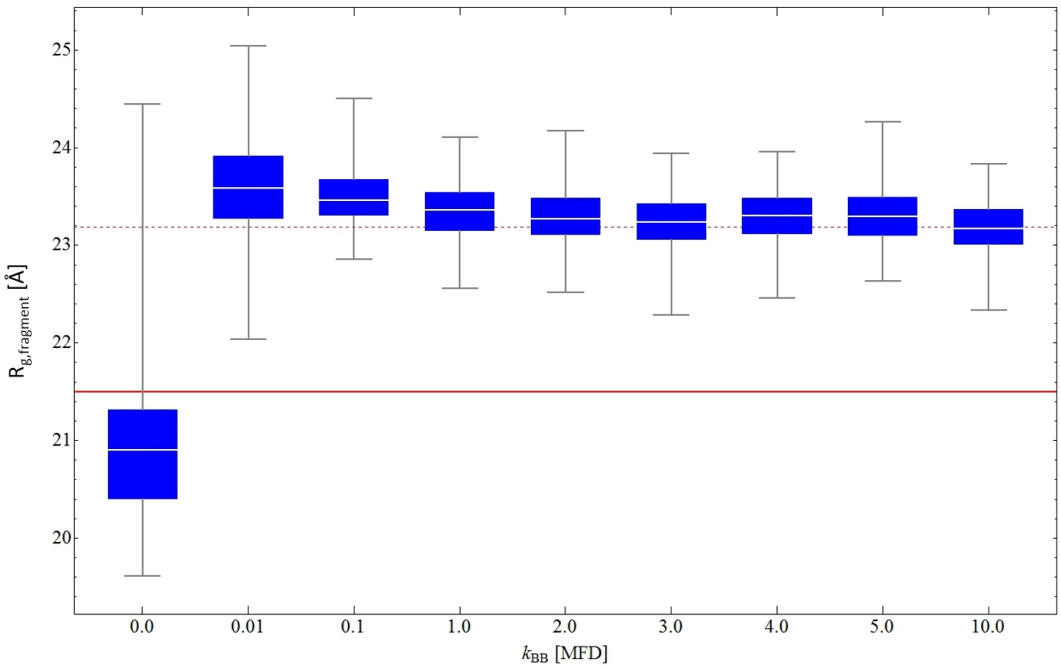


Abbildung 7.6: *Boxplot des fragmentbasierten Gyrationsradius $R_{g,fragment}$ in Abhängigkeit der Stärke der Kraftkonstante k_{BB}.*

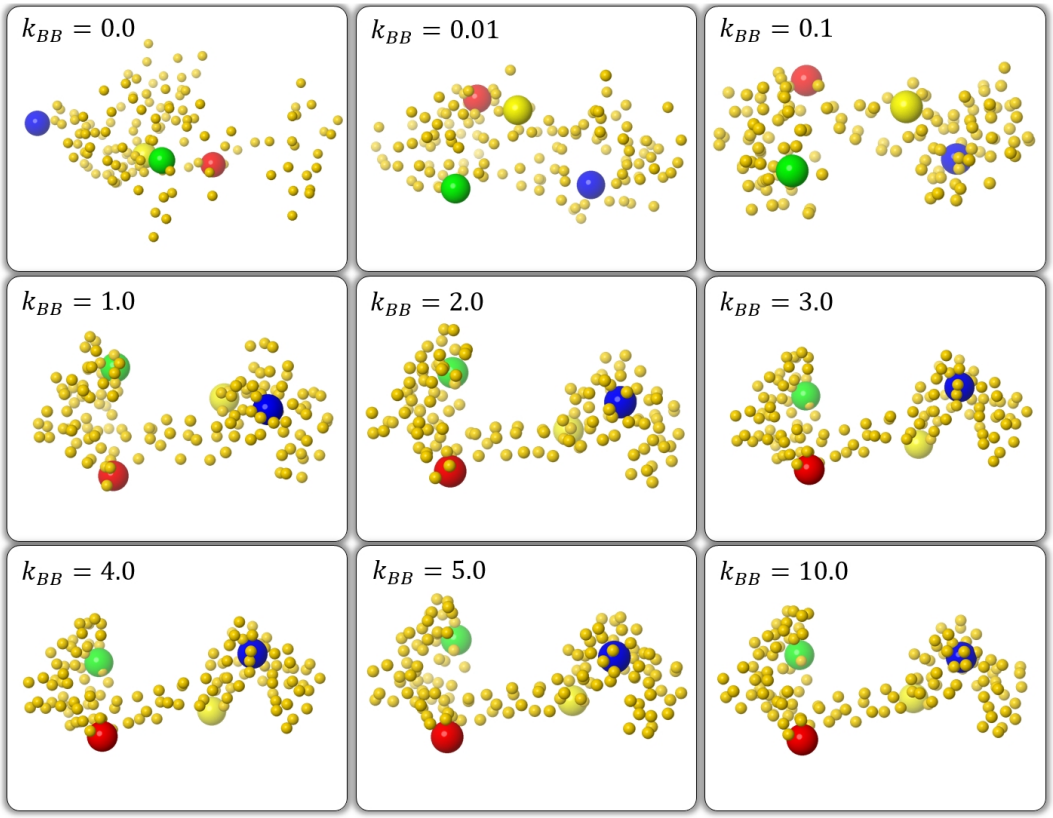


Abbildung 7.7: *Bilder der Simulationsbox für jeweils den letzten Iterationsschritt ($t = 1.51\ \mu s$) der angegebenen Kraftkonstante k_{BB}.*

Wird die Kraftkonstante der harmonischen Potentiale auf $k_{BB} = 0.0$ gesetzt, ist der Gyrationsradius mit einem Wert von $R_{g,fragment} = 21.03$ Å kleiner als der Gyrationsradius mit einer hohen Kraftkonstante von $k_{BB} = 10.0$ und einem Wert von $R_{g,fragment} = 23.18$ Å. Der Grund hierfür ist, dass die zentrale α-Helix des Proteins kollabiert ist und sich eine globuläre Proteinkonformation herausbildet (Abbildung 7.7, $k_{BB} = 0.0$). Anhand der Differenz der mittleren Distanz des Sondenpaares 1 - 3 (Abbildung 7.8) gegenüber der Distanz in der Kristallstruktur Δd_{13} lässt sich die Konformation der α-Helix überprüfen. Für $k_{BB} = 0.0$ beträgt $\Delta d_{13} = 13.88$ Å, was auf eine kollabierte Konformation des Calmodulins schließen lässt. Ab einer Kraftkonstante von $k_{BB} = 0.1$ ist $\Delta d_{13} = 0.06$ Å. Dies zeigt, dass ab diesem Wert die mittlere Konformation des Proteins erhalten bleibt und die zentrale α-Helix ausgeprägt ist. Zusätzlich ist anhand von Abbildung 7.6 und 7.8 zu erkennen, dass mittels einer größer gewählten Kraftkonstante die Spannweite der Distanzverteilung sowie die Streuung des fragmentbasierten Gyrationsradius gegenüber der Kristallstruktur verringerbar ist. Somit ist die Flexibilität des Proteinrückgrates über die Stärke der Kraftkonstante k_{BB} steuerbar. Deutlich zu erkennen ist die fortschreitende Konservierung der Struktur durch eine höhere Kraftkonstante in Abbildung 7.7. Die Strukturen für eine Kraftkonstante von $k_{BB} = 0.0$ und $k_{BB} = 0.01$ wirken noch diffus und weichen deutlich von der Tertiärstruktur ab. Für $k_{BB} \geq 0.1$ ist die hantelförmige Konformation des Calmodulins erkennbar und die zentrale α-Helix ausgeprägt. Ab $k_{BB} \geq 1.0$ ist visuell kein Unterschied zur Kristallstruktur mehr erfassbar.

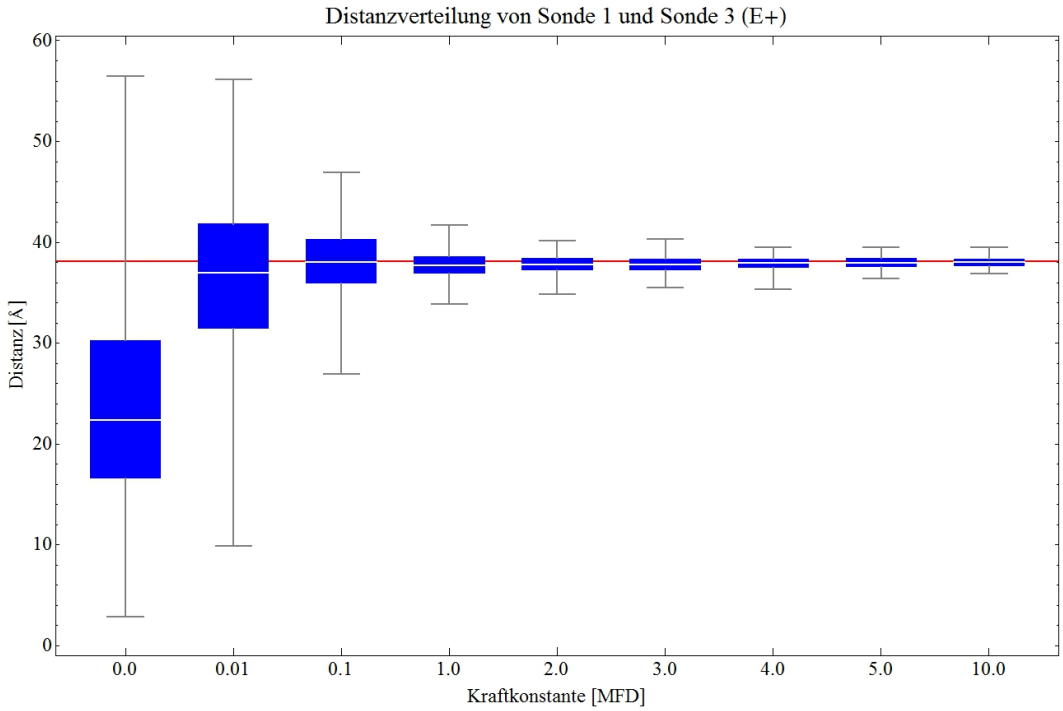


Abbildung 7.8: *Boxplot der Distanzverteilungen von Sonde 1 und Sonde 3.*

Im Anschluss an die Betrachtung der Auswirkungen der Änderung der Kraftkonstante k_{BB} soll der Einfluss der Länge der aktivierten Potentiale d_{BB} bei konstanter Kraftkonstante mit einem Wert von $k_{BB} = 2.0$ untersucht werden. Abbildung 7.9 zeigt die Entwicklung des fragmentbasierten Gyrationsradius in Abhängigkeit der aktivierten Potentiale, wobei die Ergebnisse des sukzessiven Hinzufügens von kurzreichweitigen Potentialen rot, von langreichweitigen Potentialen blau und die Kombination aus kurz- und langreichweitigen Potentialen grün dargestellt ist. Deutlich zu erkennen ist, dass alle Verläufe ab einer Anzahl von mehr als 5000 Potentialen in den gemittelten Gyrationsradius von 22.8 Å übergehen. Das beste Ergebnis zeigen die langreichweitigen Potentiale. Ab 2628 Potentialen beträgt die Spannweite 1.4 Å. Auch der Mittelwert des fragmentbasierten Gyrationsradius $R_{g,fragment}$ stimmt mit 23.2 Å gut mit dem Wert der höchstkonserviertesten Simulation mit 22.8 Å überein. Die leichte Erhöhung entsteht wahrscheinlich dadurch, dass das Protein aufgrund der ausschließlich langreichweitigen Wechselwirkungen auseinander gedrückt wird. Das schlechteste Ergebnis zeigen die kurzreichweitigen Potentiale. Ab 3968 Potentialen liegt der Gyrationsradius bei 23.3 Å, also leicht höher als der Wert der langen Potentiale. Auch wird das Protein wieder

minimal auseinander gedrückt. Jedoch zeigt die Spannweite mit 4.4 Å eine geringe Konservierung des Proteins. Erst ab 6468 Potentialen erreicht die Spannweite einen Wert von 1.8 Å. Die kombinierten Potentiale zeigen einen Kurvenverlauf, der eine Mischung aus den kurz- und langreichweitigen Potentialen darstellt.

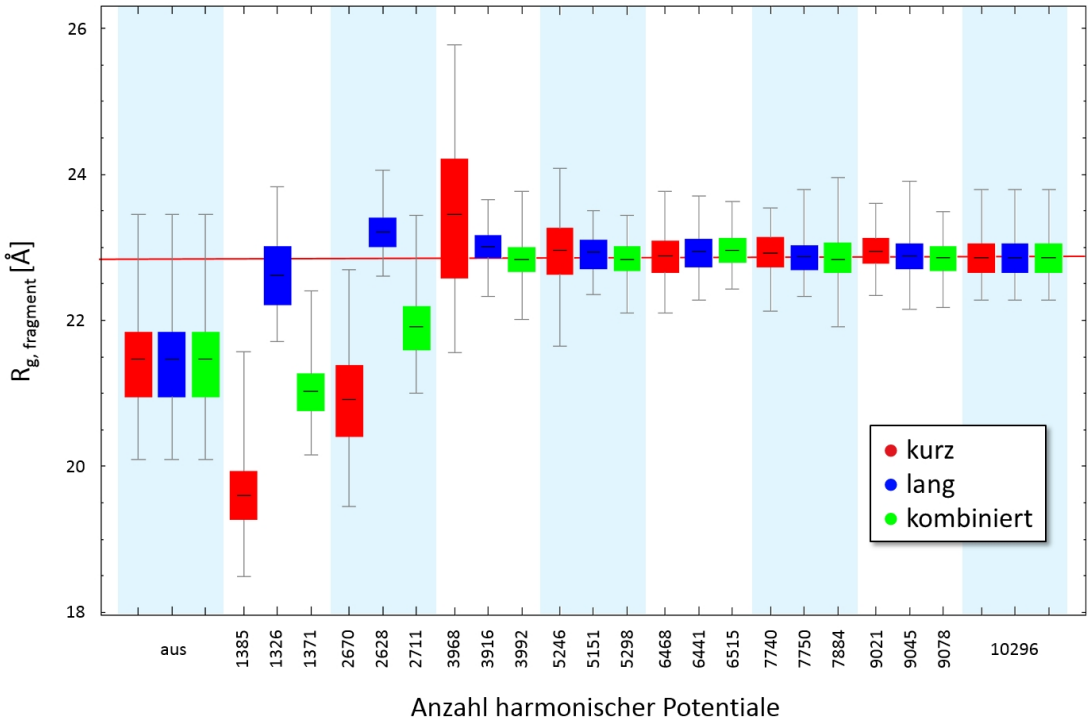


Abbildung 7.9: *Boxplot des fragmentbasierten Gyrationsradius in Abhängigkeit von den aktivierten Distanzlängen der jeweiligen Rückgratpotentiale (rot: Kurzreichweitige Potentiale; Blau: Langreichweitige Potentiale; Grün: Kombination aus kurz- und langreichweitigen Potentialen).*

Abbildung 7.10 zeigt den Einfluss der aktivierten Distanzlängen der Rückgratpotentiale auf den Abstand zwischen den Sonden 1 und 3. Die Betrachtung dieses Sondenpaares ist besonders interessant, da die Sonden in den gegenüberliegenden EF-Hand-Motiven des Calmodulins lokalisiert sind. Die Distanz des Sondenpaares ist somit direkt abhängig von der Konformation der mittleren α-Helix. Der Mittelwert der Distanz zwischen den Sonden 1 und 3 für die am höchsten konservierte Simulation liegt bei 36.3 Å und die Spannweite bei 4.9 Å. Die beste Leistung zeigen auch bei den Distanzen die langreichweitigen Potentiale. Bei 2628 akti-

vierten Potentialen ist der Mittelwert 38.0 Å und die Spannweite beträgt 14.4 Å. Die schlechtesten Konservierungseigenschaften zeigen die kombinierten Potentiale. Der Mittelwert liegt zwar bei 2711 aktivierten Potentialen mit 35.7 Å sehr nahe bei dem Wert der am höchsten konservierten Simulation, jedoch die Spannweite ist mit 33.2 Å sehr hoch. Die kurzreichweitigen Potentiale bilden auch schon bei 2670 Potentialen eine Distanz von 37.7 Å aus, aber die Spannweite der Werte liegt ungefähr 20% höher als bei den langreichweitigen Potentialen. Ab ungefähr 6400 zeigen alle Potentialsysteme das Verhalten wie die am höchsten konservierte Simulation.

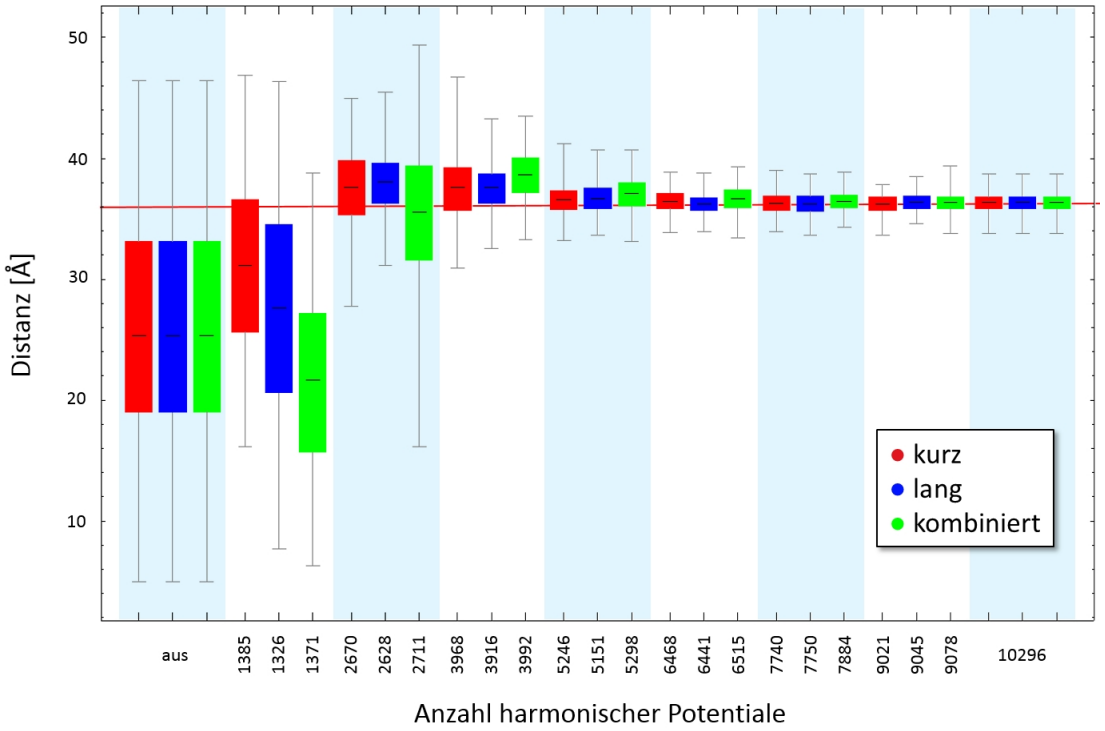


Abbildung 7.10: *Boxplot der Distanzverteilungen von Sonde 1 und Sonde 3.*

Betrachtet man den Winkel zwischen den Sonden 1-3-4, ergibt sich, dass die Stabilisierung der Winkel aufwendiger ist als die der Distanzen. Der Mittelwert der am höchsten konservierten Simulation ist 113.0° und die Spannweite beträgt 31.6°. Auch hier zeigen die langreichweitigen Potentiale wieder die beste Leistung mit einem Mittelwert von 103.2° und einer Spannweite von 82.7° bei 2628 aktivierten Potentialen. Aber wie man aus der Spannweite ableiten kann, sind

die Änderungen des Winkels während der Simulation erheblich. Die zweitbeste Konservierung zeigen die kurzreichweitigen Potentiale. Zwar zeigt sich auch hier schon bei 2670 aktivierten Potentialen ein Mittelwert von 106.1°, jedoch ist dieses Ergebnis mit einer Spannweite von 110.2° verbunden. Die kombinierten Potentiale erreichen erst mit 3992 die Werte der Vorgänger (Mittelwert von 112.1° mit einer Spannweite von 47.8°). Erst ab 6400 geht das Verhalten aller Systeme in das Verhalten der am höchsten konservierten Simulation über.

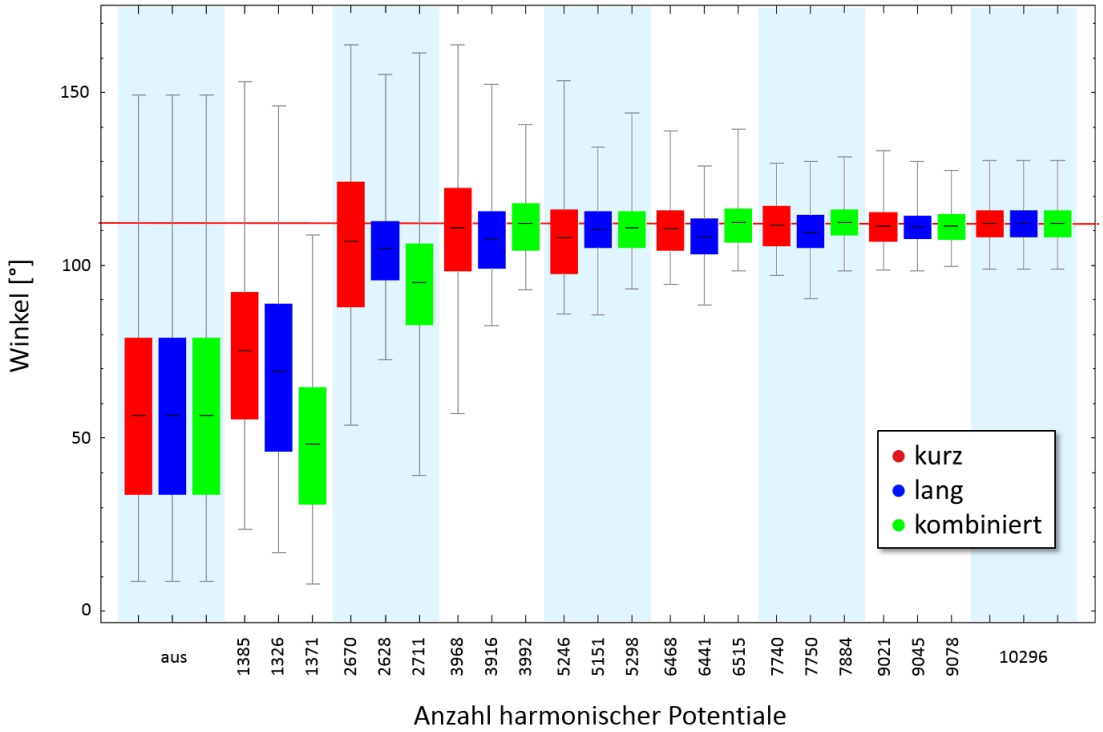


Abbildung 7.11: *Boxplot der Winkelverteilungen der Sonden 1-3-4.*

In Abbildung 7.12 sind Ansichten für kurze, lange und kombinierte Potentiale des Calmodulins nach einer Simulationszeit von 1.5 μs dargestellt. Die Hantelstruktur des Calmodulins ist für kombinierte und lange Potentiale ab ungefähr 2600 aktivierten Potentialen erkennbar, wobei die zentrale α-Helix bei den langen Potentialen besser ausgeprägt ist. Alle Systeme zeigen eine deutliche Ausprägung der zentralen α-Helix mit ca. 3900 aktivierten Potentialen.

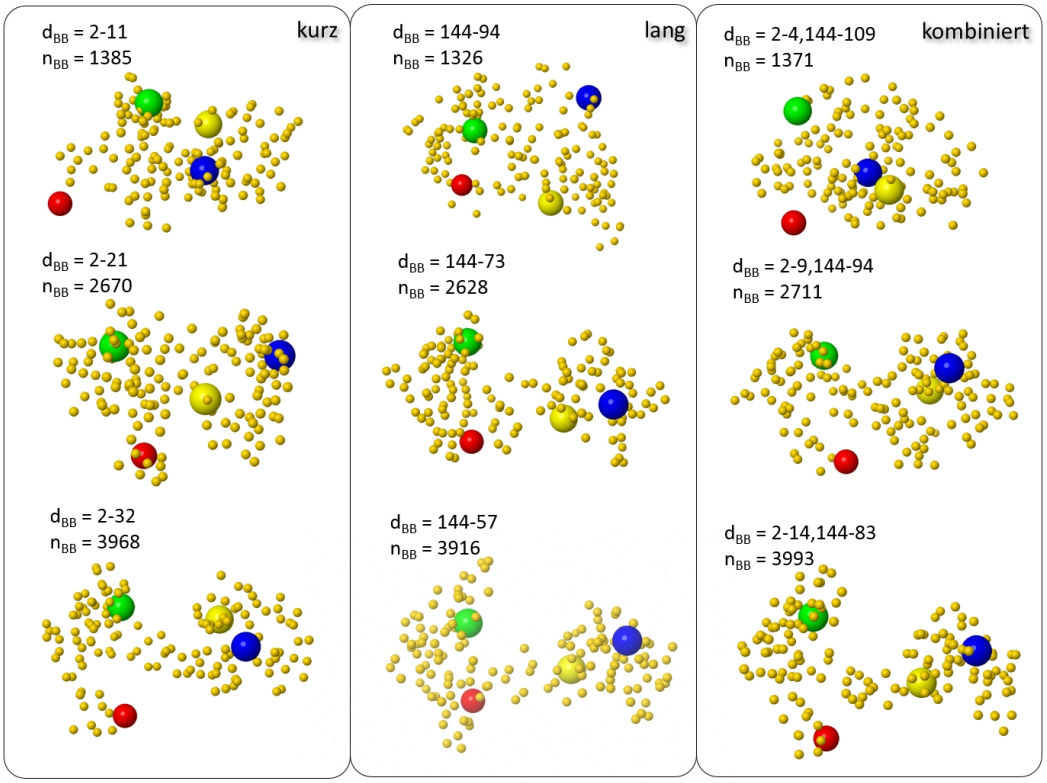


Abbildung 7.12: *Ansicht des Calmodulins für alle Potentialsysteme für den letzten Iterationsschritt der jeweiligen Simulation.*

Insgesamt konnte gezeigt werden, dass das hier erarbeitete Stabilisierungschema der Proteingeometrie erfolgreich die Konformation aufrechterhält. Außerdem besteht die Möglichkeit, dass die Flexibilität der Stabilisation über die Kraftkonstante k_{BB} frei einstellbar ist. Als Daumenregel kann für durchschnittlich große Proteine (ca. 316 Aminosäuren[45]) eine Kraftkonstante k_{BB} mit einem Wert von 1.0 gewählt werden, wenn alle Potentiale aktiviert sind. Jedoch ist es empfehlenswert, eine Simulationsreihe zur Bestimmung einer angemessenen Stärke der Kraftkonstante k_{BB} in Abhängigkeit der individuellen Simulationsparameter durchzuführen. Zusätzlich wird ersichtlich, dass insbesondere die langreichweitigen Potentiale für die Strukturkonservierung am wichtigsten sind. All dies ermöglicht die Untersuchung von Proteinsystemen mit einer hohen Strukturkonservierung, aber auch von Systemen mit hoher Flexibilität.

7.2 Langzeitanalyse der Stabilität von heteromeren Proteinen

Hämoglobin ist ein eisenhaltiger Proteinkomplex, bestehend aus je zwei Untereinheiten Hb α und Hb β. In dieser Studie soll anhand des Proteins Hämoglobin (PDB ID: 2HHB[198]) die Stabilität der Quartärstruktur betrachtet werden, wobei nur die Tertiärstruktur der einzelnen Untereinheiten stabilisiert wird.

Methode

Die Simulationsbox besteht aus 10^5 Wasserfragmenten und einem zentral platzierten Hämoglobinmolekül. Die Seitenlänge der quadratischen Simulationsbox beträgt 230.9 Å. Die periodischen Randbedingungen sind in x-, y- und z-Raumrichtung aktiviert. Die Proteinrückgratsegemente sind entsprechend ihrer Zugehörigkeit zu der jeweiligen Proteinuntereinheit nummeriert, sodass die Stabilisierungspotentiale nicht übergreifend auf unterschiedliche Proteinuntereinheiten wirken. Die Bewegung der einzelnen Untereinheiten unterliegt nur der Simulationsdynamik. Die Kraftkonstante der Proteinrückgratstabilisierung k_{BB} beträgt den Wert 2.0. Die Simulationszeit umfasst 10^6 Dynamikschritte, was einer Zeitperiode von 76.5 μs entspricht. Die Parameter der elektrostatischen Kraft sind in Tabelle 7.4 dargestellt.

Tabelle 7.4: *Simulationsparameter der elektrostatischen Kraft. In MFD-Einheiten, falls nicht anders angegeben .*

Parameter	Wert
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5

Abbildung 7.13 zeigt links die Quartärstruktur des Hämoglobins und rechts

die Verteilung der Sondenfragmente (rot, grün, blau, gelb) innerhalb des Proteinrückgrates (α-C-Fragmente: gold) des Hämoglobinmoleküls. Die Fragmente der Seitenketten sind ausgeblendet.

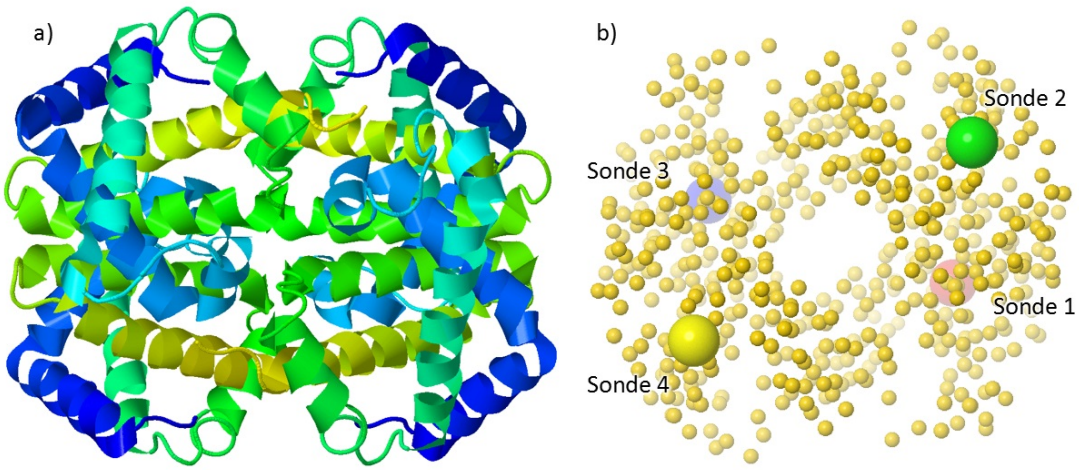


Abbildung 7.13: *a) Darstellung der Quartärtstruktur des Hämoglobins. b) Konfiguration der Sondenfragmente innerhalb des Proteinrückgrates.*

Ergebnis

Eine Betrachtung des Gyrationsradius des simulierten Hämoglobinmoleküls zeigt, dass der durchschnittliche Gyrationsradius in der Simulation $R_{g,fragment} = 28.5 \pm 0.5$ Å beträgt. In Abbildung 7.14 ist der zeitliche Verlauf des ermittelten fragmentbasierten Gyrationsradius dargestellt (fehlende Werte sind gefilterte Artefakte innerhalb der Berechnung). Der experimentell bestimmte Gyrationsradius hat einen Wert von $R_g = 29 \pm 2$ Å[199]. Daraus folgt, dass die Ausdehnung des Hämoglobins im Mittel erhalten bleibt. Die einzelnen Untereinheiten des Proteins driften somit nicht auseinander, sondern behalten im Durchschnitt die in der Quartärstruktur vorgegebene globuläre Organisation bei.

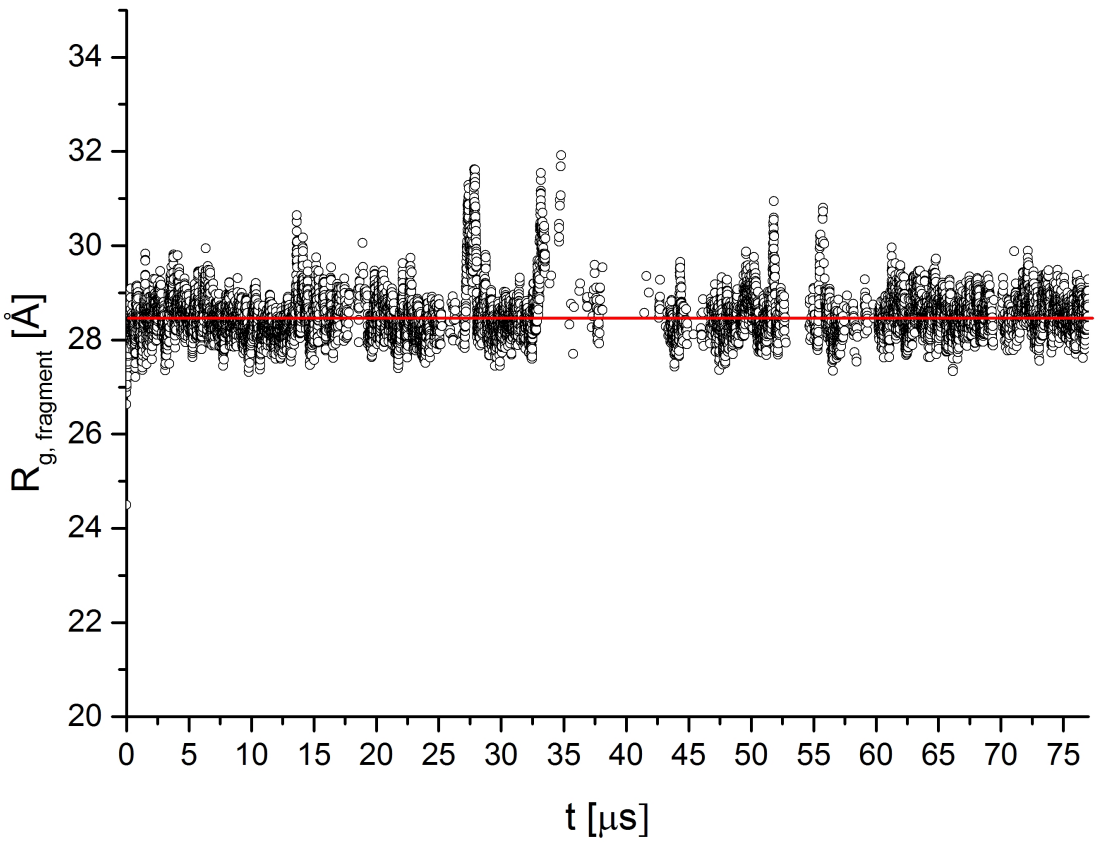


Abbildung 7.14: *Zeitlicher Verlauf des Gyrationsradius des Hämoglobinmoleküls. Die rote Linie beschreibt den Mittelwert mit* $R_{g,fragment} = 28.5$ *Å.*

Das große Diagramm in Abbildung 7.15 skizziert exemplarisch den Verlauf der Distanz zweier Sondenpaare (1 und 2, 1 und 3). Die Kurve von Sonde 1 und 2 spiegelt den Verlauf der meisten Sondenpaare wieder. Die Distanz in der Kristallstruktur beträgt für dieses Sondenpaar 41.8 Å. Die durchschnittliche Distanz zwischen den beiden Sonden innerhalb der Simulation beträgt 47.9 ± 5.3 Å. Die Standardabweichung der Distanz, wenn man den Wert der Kristallstruktur als Mittelwert wählt, beträgt 8.1 Å. Die Distanz hat sich somit im Mittel um 6.1 Å gegenüber der Kristallstruktur vergrößert.

Der Verlauf der Distanz des Sondenpaares 1 und 3 hat einen durchschnittlichen Wert von 44.8±8.8 Å. Die Distanz des Sondenpaares innerhalb der Kristallstruktur beträgt 35.05 Å. Demnach hat sich die mittlere Distanz um 9.78 Å vergrößert. Der Verlauf der Distanzen dieses Sondenpaares ist jedoch durch zwei starke Abwei-

chungen bei ungefähr $t = 27.0\ \mu s$ und $t = 34.0\ \mu s$ mit einer maximalen Distanz von 83.0 Å gekennzeichnet.

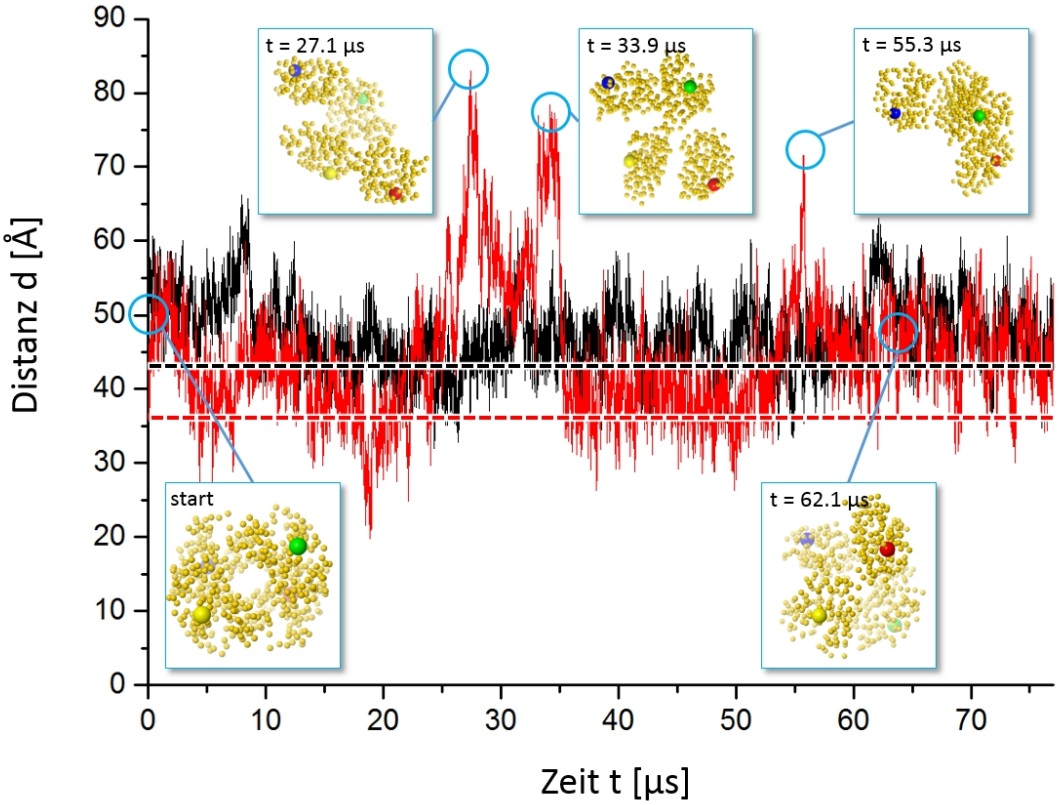


Abbildung 7.15: *Zeitliche Auflösung der Distanzen zwischen den Sondenfragmenten 1 - 2 und 1 - 3 sowie Ansichten des Hämoglobins zu hervorhebenswerten Zeitpunkten. Die gestrichelten Linien zeigen die Distanzen innerhalb der Kristallstruktur.*

In Abbildung 7.16 ist der zeitliche Verlauf der Winkelveränderung zwischen den Sondenfragmenten 2 - 3 - 4 und 1 - 4 - 3 dargestellt. Betrachtet man den Winkel zwischen Sonde 2 - 3 - 4, so beträgt dieser in der Kristallstruktur 52.9°. In der Simulation zeigt sich ein Winkel mit einem Wert von $63.0° \pm 10.0°$. Dies entspricht einer durchschnittlichen Vergrößerung des Winkels um 10.2°.

Die Ergebnisse des Winkels der Sonden 1 - 4 - 3 betragen in der Kristallstruktur 42.9°. Der simulierte Winkel hat eine Größe von $56.3° \pm 16.9°$. Zudem vermittelt der Winkel das gleiche Bild, wie die Distanzen der Sonden 1 und 2. Bei ungefähr $t = 27.0\ \mu s$ und $t = 34.0\ \mu s$ sind zwei Spitzen mit einem maximalen Wert von

163.9° im Verlauf zu erkennen.

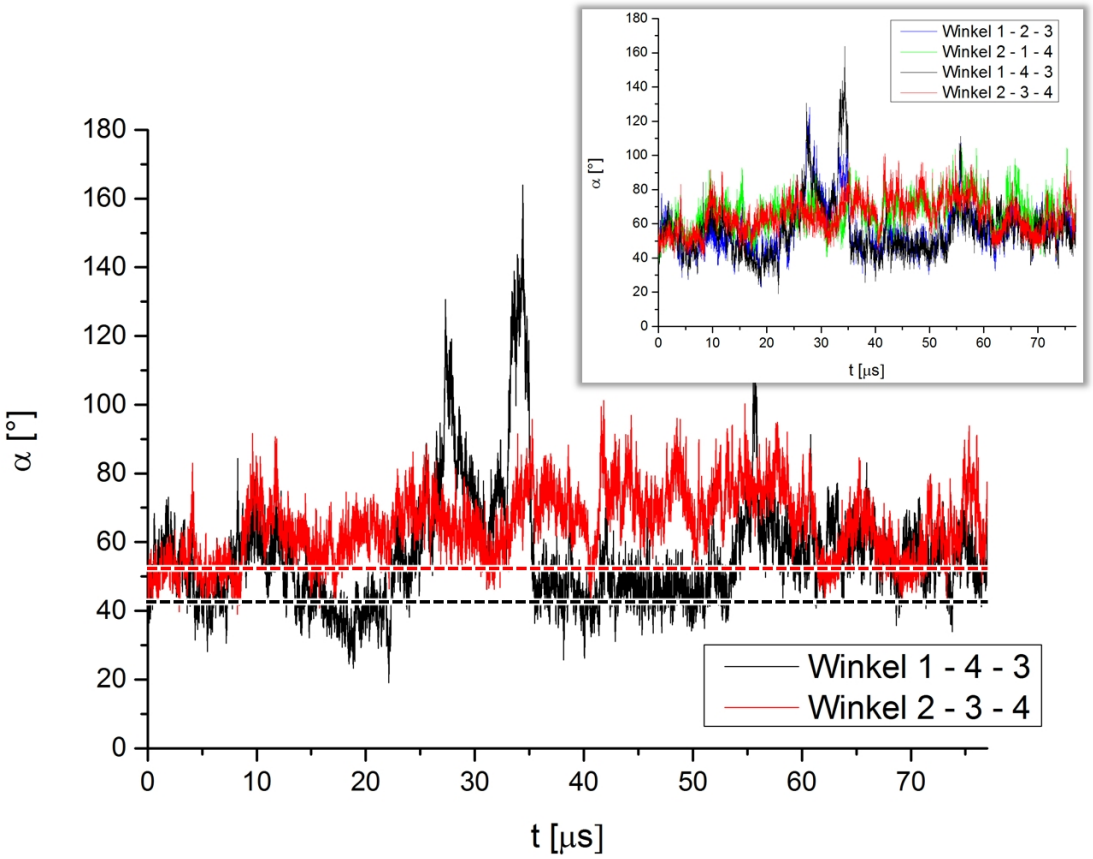


Abbildung 7.16: *Zeitlicher Verlauf der Winkel der Sondenfragmente 2 - 3 - 4 (schwarz) und 1 - 4 - 3 (rot). Die gestrichelten Linien zeigen die Winkel innerhalb der Kristallstruktur.*

Abbildung 7.17 skizziert verschiedene Ansichten des Hämoglobinmoleküls zu unterschiedlichen Zeiten. Bis zum Zeitpunkt $t = 34.4$ μs besitzt das Hämoglobinmolekül eine globuläre Struktur. Bei $t = 34.4$ μs erkennt man deutlich eine planare Struktur. Die Sondenfragmente liegen fast genau in einer Ebene. Dies spiegelt sich auch in den vorher beschriebenen Distanzen der Sonden 1 - 3 sowie dem Winkel der Sonden 1 - 4 - 3 wieder. Der Winkel der Sonden 1 - 4 - 3 näherte sich mit 163.9° fast 180.0° an, welches die planare Struktur der Hämoglobinuntereinheiten wiedergibt.

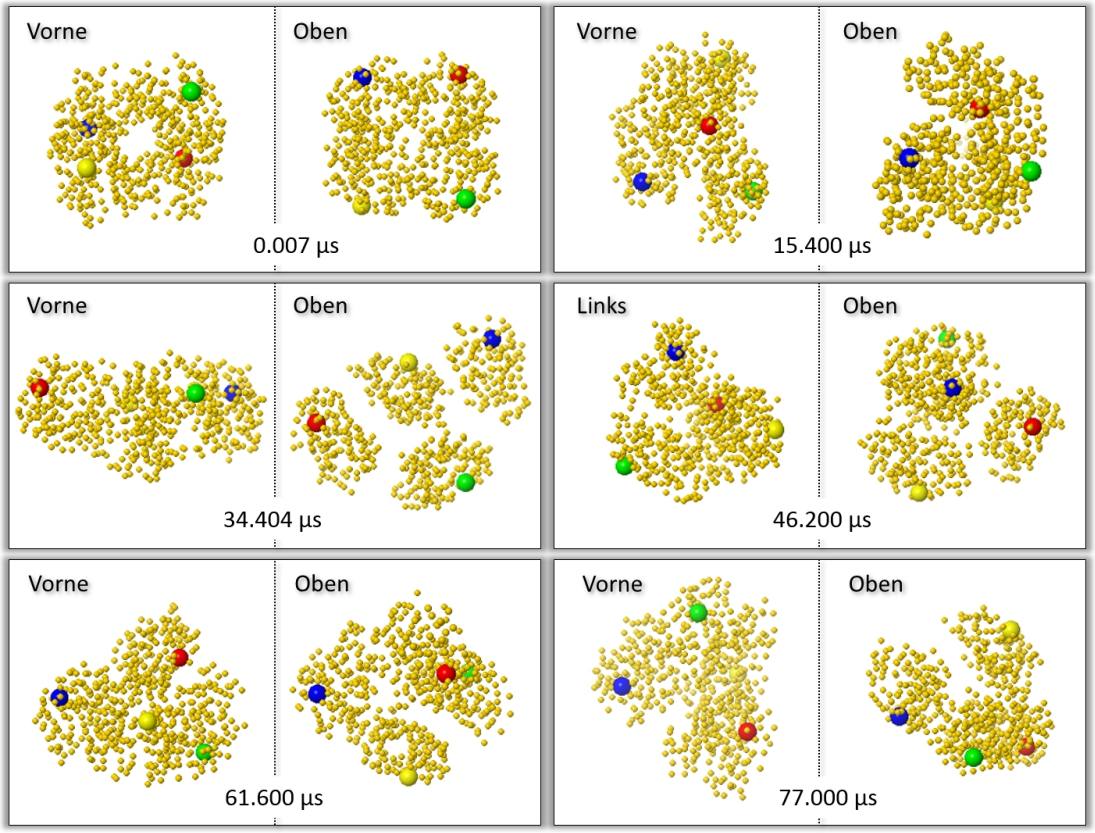


Abbildung 7.17: *Ansicht des Hämoglobinmoleküls während der Simulation zu unterschiedlichen Zeiten und Perspektiven.*

Die Simulation des Hämoglobinmoleküls hat gezeigt, dass die bevorzugte Konformation des Moleküls eine globuläre Struktur ist. Das nur kurze Auftreten einer planaren Ausrichtung der Untereinheiten zueinander stellt eine geringere Stabilität dieser Struktur gegenüber der globulären Struktur dar. Die Untereinheiten standen während der kompletten Simulation im Kontakt miteinander. Experimentelle Daten bezüglich der Dynamik von Proteinen in realen Systemen sind auf Grund der Komplexität des Fragestellung nicht verfügbar.

7.3 Studie zur Interaktion von Zinkricinoleat mit histidinhaltigen Peptiden und Proteinen

Das Patent DE102011101880 A1[69] beschreibt, dass Zinkricinoleat ($Zn(Ri)_2$) fähig ist als neues Auftrennungsmedium gegenüber konventionellen Medien auf Basis von Metallchelaten zur Auftrennung von Proteinen und Enzymen mit einem Polyhistidin-Tag zu fungieren. Dennoch ist der Mechanismus der Bindung von His-Tag Proteinen an eine Zinkricinoleatschicht noch nicht aufgeklärt und daher Bestandteil dieser Studie. Jede nachfolgende Simulation wird bei 310 K durchgeführt. Die periodischen Randbedingungen sind in x- und y-Raumrichtung aktiviert. In z-Richtung sind die periodischen Randbedingungen deaktiviert, um eine Diffusion der $Zn(Ri)_2$-Moleküle über die Boxgrenze hinaus zu verhindern und somit die Zerstörung der gebildeten Schicht zu vermeiden. Vor dem Start der jeweiligen Simulation werden 300 Minimierungsschritte durchgeführt. Der angewendete pH-Wert beträgt 7.4 für alle Peptide und Proteine. Die Parameter der elektrostatischen Kraft sind in Tabelle 7.5 dargestellt.

Tabelle 7.5: *Simulationsparameter der elektrostatischen Kraft.*

Parameter	Wert [MFD]
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5

Das Fragmentierungsschema des $Zn(Ri)_2$ Moleküles ist in Abbildung 7.18 skizziert.

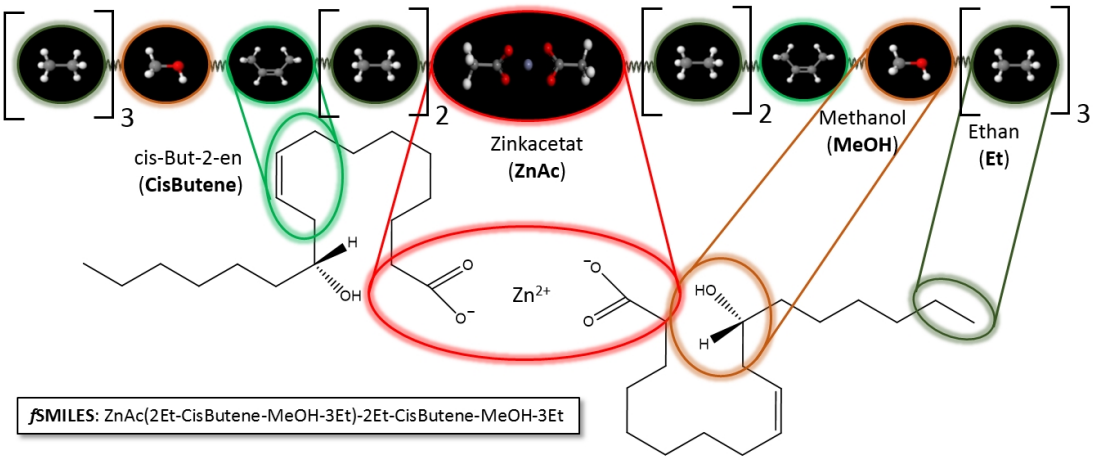


Abbildung 7.18: *Fragmentierungsschema, Struktur und fSMILES des Zn(Ri)$_2$-Moleküls.*

Die Fragmente, die die Ricinolsäure repräsentieren, sind grün und die Methanolfragmente der Hydroxygruppe der Ricinolsäure braun gefärbt. Die Zinkacetatkopfgruppe ist rot eingefärbt.

Das Fragmentierungsschema der Aminosäure Histidin ist in Abbildung 7.19 dargestellt. Die Imidazolgruppe des Histidins ist gelb gefärbt. Alle anderen Fragmente erhalten die Farbe blau.

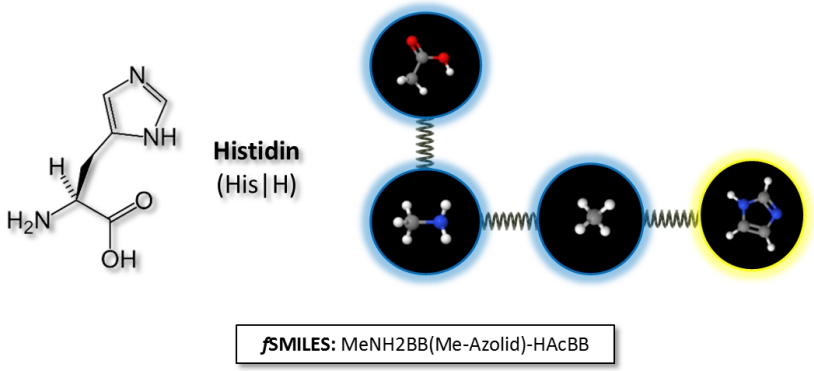


Abbildung 7.19: *Fragmentierungsschema der Aminosäure Histidin.*

7.3.1 Interaktion von Aminosäuren mit einer $Zn(Ri)_2$-Schicht

Methode

Zur Darstellung der Wechselwirkung eines einzelnen Histidinmoleküls mit einer $Zn(Ri)_2$-Schicht wird eine kubische Simulationsbox mit der Kantenlänge von $32{\cdot}32{\cdot}32$ nm^3 erstellt. Die Box enthält $2{\cdot}10^5$ Wassermoleküle und ein einzelnes Histidinmolekül über einer $Zn(Ri)_2$-Schicht. Die $Zn(Ri)_2$-Schicht besteht aus 1600 Molekülen, welche in einer xy-Schicht in der Mitte der Box platziert sind. Die periodischen Randbedingungen sind in diesem Fall in allen Raumrichtungen aktiviert. Die Simulationszeit beträgt 1.7 μs ($2{\cdot}10^4$ Iterationen).

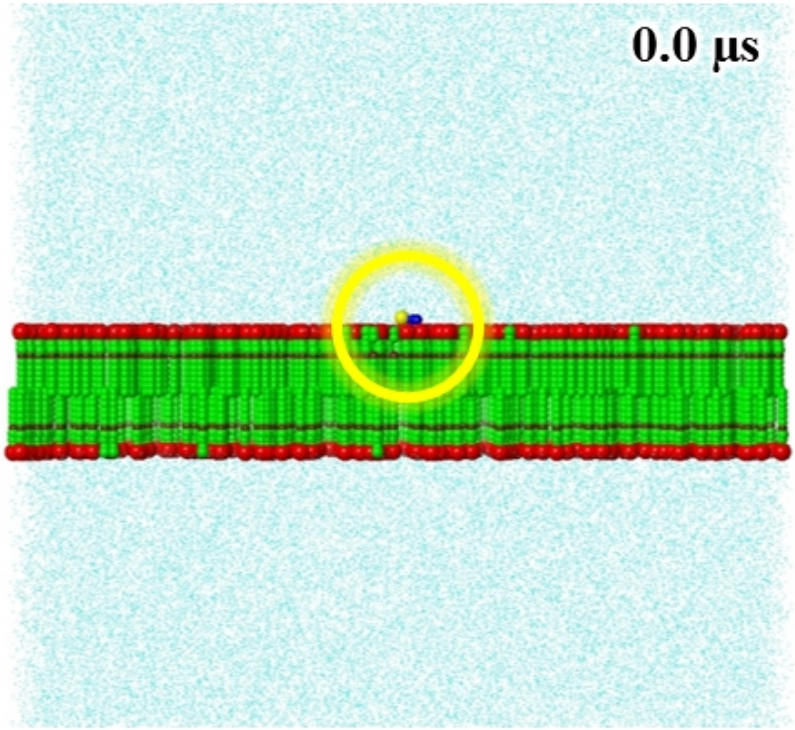


Abbildung 7.20: *Startgeometrie der Simulationen mit einem Histidin.*

Für die Untersuchung der Bindungsselektivität von Histidin gegenüber anderen Aminosäuren an einer $Zn(Ri)_2$-Schicht werden Simulationsboxen mit 1 : 1 Mischungen von Histidin und einer anderen Aminosäure erstellt. Die Simulationsboxen haben eine Größe von $28{\cdot}28{\cdot}21$ nm^3. Die Höhe der Box ist reduziert, um

eine größere Oberfläche der $Zn(Ri)_2$-Schicht zu erreichen. Die Box beinhaltet 10^5 Wassermoleküle und 1000 $Zn(Ri)_2$-Moleküle, welche in einer xy-Schicht am Boden der Box platziert werden. Je 100 Moleküle der jeweiligen Aminosäure werden eingesetzt. Über der $Zn(Ri)_2$-Schicht befindet sich eine Schicht aus 40 % aller Wassermoleküle. Die Aminosäuren werden zufällig in der obersten Schicht mit den restlichen 60 % der Wassermoleküle positioniert. Neben den Minimierungsschritten zu Beginn der Simulation werden 440 ns simuliert, wobei die Positionen der Aminosäuren fixiert sind, damit sich die $Zn(Ri)_2$-Schicht ungestört ausbilden kann. Die gesamte Simulationszeit beträgt 3.5 μs (40000 Iterationen). Die verwendeten Aminosäuren neben Histidin sind Phenylalanin, Alanin, Leucin, Isoleucin, Prolin und Tryptophan. Abbildung 7.21 A zeigt die Anordnung der einzelnen Schichten und Abbildung 7.21 B die resultierenden Simulationsboxen.

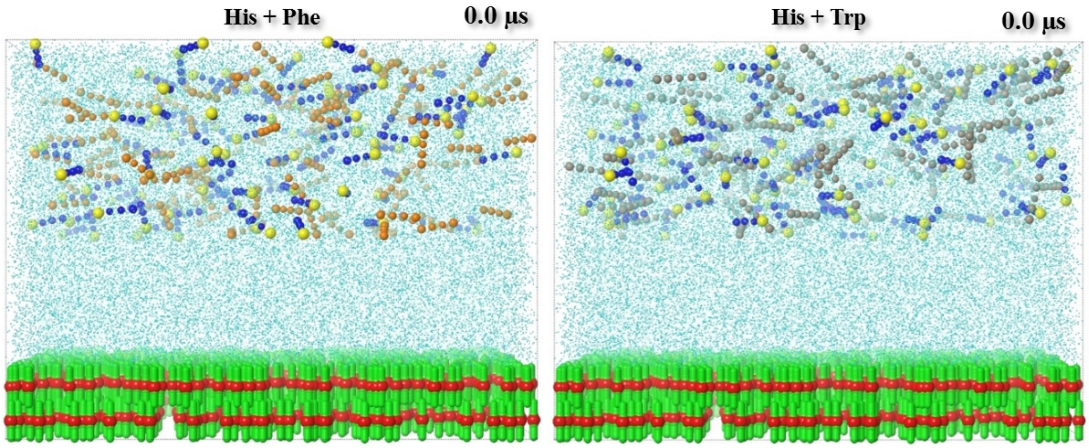


Abbildung 7.21: *Startgeometrie der Simulationen mit 100 Histidinmolekülen und 100 konkurrierenden Aminosäuremolekülen.*

Ergebnis

Während der gesamten Simulationszeit von 1.7 μs diffundiert das einzelne Histidinmolekül nahe an der Oberfläche der $Zn(Ri)_2$-Schicht. Häufig interagiert das Imidazolfragment (gelb) direkt mit den Zinkacetatkopfgruppen (rot) der $Zn(Ri)_2$-Oberfläche. Dieser Kontakt ist jedoch nicht permanent. Hauptsächlich steht das Histidin in Kontakt mit der $Zn(Ri)_2$-Schicht und diffundiert nicht frei in der Simulationsbox. Dies unterstützt die Vermutung der Interaktion zwischen der stickstoff-

haltigen Gruppe des Imidazols und dem Zinkricinoleat.

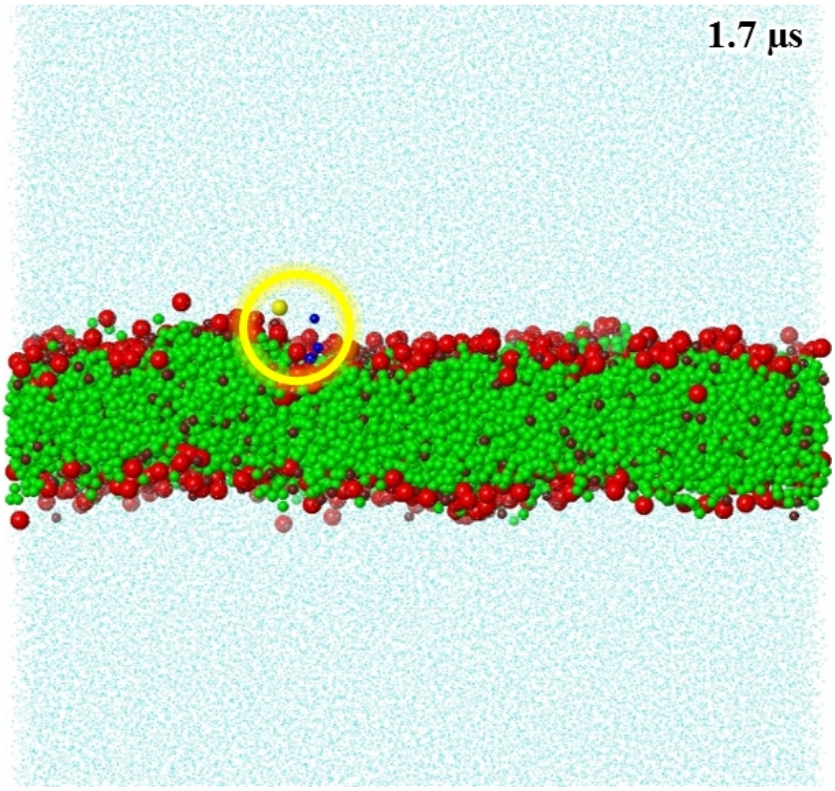


Abbildung 7.22: *Querschnitt der Simulationsbox mit einem Histidinmolekül.*

Die Ergebnisse der Aminosäuregemische sind für Histidin - Phenylalanin und Histidin - Tryptophan in Abbildung 7.23 gezeigt. Nach einer Simulationszeit von 3.5 μs ist keine besonders starke Wechselwirkung zwischen Phenylalanin sowie Histidin mit der Oberfläche der $Zn(Ri)_2$-Schicht zu erkennen. Nur geringfügig mehr Histidinmoleküle stehen in Kontakt mit der Oberfläche. Die meisten Aminosäuren diffundieren frei in der Simulationsbox. Ein interessanter Aspekt ist, dass die Phenylalaninmoleküle mit Kontakt zu der $Zn(Ri)_2$-Schicht in der hydrophoben Schicht anzutreffen sind und nicht bei den hydrophilen Zinkacetatkopfgruppen. Für die Mischung Histidin - Tryptophan zeigt sich das selbe Bild. Nur geringfügig mehr Histidinmoleküle stehen in Kontakt mit der $Zn(Ri)_2$-Schicht. Auch hier sind die Tryptophanmoleküle in den hydrophoben Anteilen der $Zn(Ri)_2$-Schicht anzutreffen. Die Ergebnisse der anderen Aminosäuren sind äquivalent und werden hier nicht dargestellt. Zusammenfassend kann gesagt werden, dass eine erhöhte Wech-

selwirkung der Histidinmoleküle speziell des Imidazolfragmentes mit der $Zn(Ri)_2$-Oberfläche verglichen mit den anderen Aminosäuren vorliegt. Dies ist an der ausschließlich an der $Zn(Ri)_2$-Oberfläche auftretenden Wechselwirkung des Histidinmoleküls erkennbar. Eine explizite Bevorzugung der Absorption von Histidin gegenüber anderen Aminosäuren konnte jedoch nicht gezeigt werden.

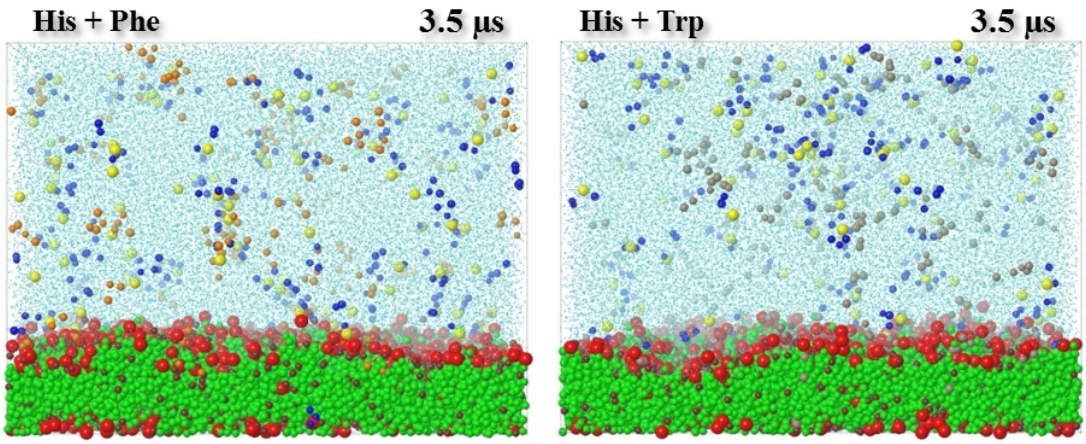


Abbildung 7.23: *Querschnitt der Simulationsbox mit 100 Histidinmolekülen und 100 Phenylalanin- bzw. Tryptophanmolekülen.*

7.3.2 Polyhistidinpeptide interagieren stärker mit einer Zinkricinoleatschicht als mit einzelnen Histidinmolekülen

Methode

Um die stärkere Wechselwirkung von $Zn(Ri)_2$-Schichten mit Polyhistidinpeptiden gegenüber einzelnen Histidinmolekülen darzustellen, wird eine quadratische Simulationsbox mit der Größe von 28·28·21 nm^3 konstruiert. Die Höhe der Box ist reduziert, um eine größere Oberfläche der $Zn(Ri)_2$-Schicht zu erreichen. Die Box beinhaltet 10^5 Wassermoleküle und 1000 $Zn(Ri)_2$-Moleküle, welche in einer xy-Schicht am Boden der Box platziert werden. Drei Simulationen werden durchgeführt, wobei 100 Histidinpeptide mit unterschiedlicher Kettenlänge von 6, 8 und 10 Histidinmolekülen eingesetzt werden. Die Peptide werden in der Startgeometrie in einer Schicht 5 nm über der $Zn(Ri)_2$-Schicht platziert. Die Simulationszeit

beträgt 4.4 μs (50000 Iterationen). Abbildung 7.24 skizziert die Startgeometrien der drei Simulationen.

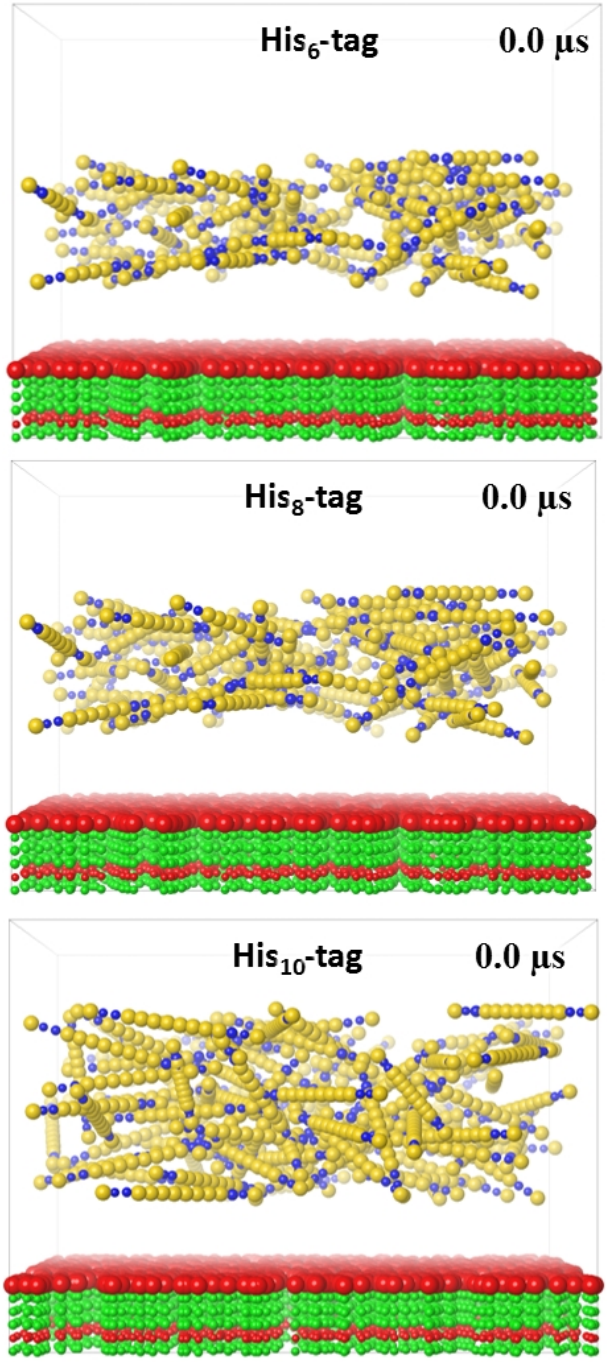


Abbildung 7.24: *Startgeometrien der Simulationen mit jeweils 100 His$_6$-, His$_8$- und His$_{10}$-Peptiden.*

Ergebnis

Abbildung 7.25 zeigt die jeweiligen Simulationsboxen der His_6-, His_8- und His_{10}-Peptide nach einer Simulationszeit von 4.4 μs. Nicht alle Histidinpeptide werden in der $Zn(Ri)_2$-Schicht absorbiert, sondern diffundieren frei in der Simulationsbox umher.

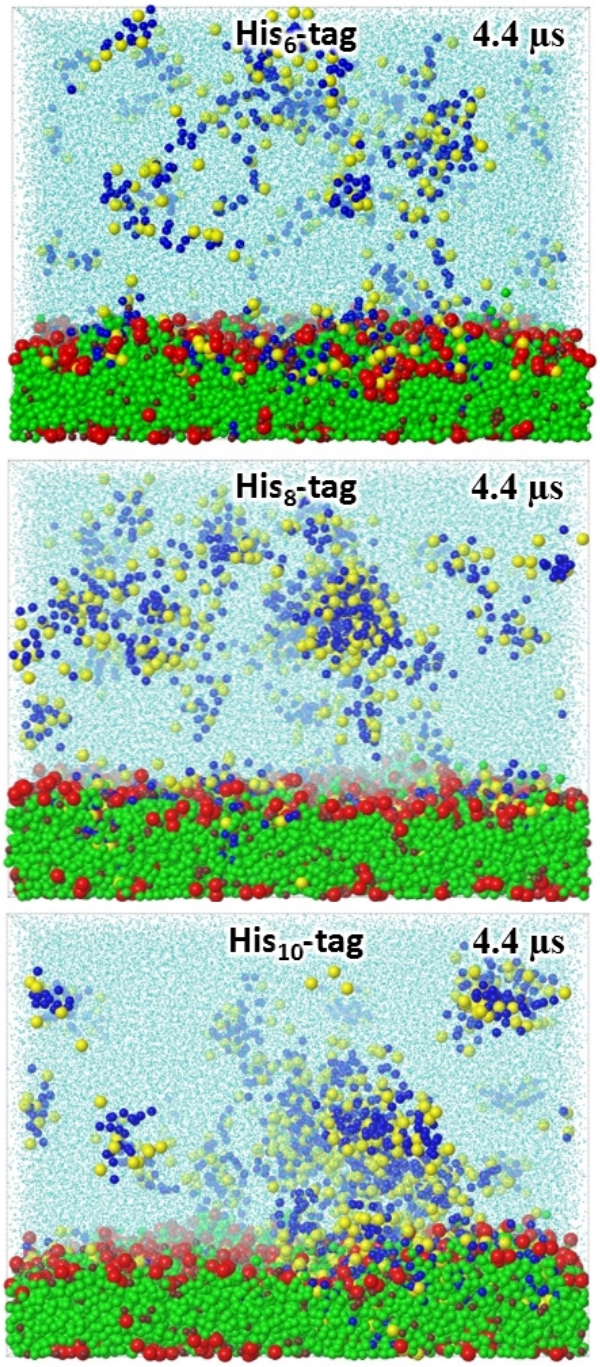


Abbildung 7.25: *Querschnitte der Simulationsboxen nach 4.4 μs Simulationszeit.*

Wie schon in den vorherigen Simulationen dargelegt, findet die Interaktion des Histidins über die Imidazolgruppe mit der Zinkacetatgruppe des $Zn(Ri)_2$ statt. Um die Anzahl der Histidinpeptide im engen Kontakt zu der $Zn(Ri)_2$-Schicht zu bestimmen, wird die Anzahl der Imidazolfragmente im Bereich des Bodens der Box bis knapp über die Zinkacetatfragmente bestimmt. Das Diagramm in Abbildung 7.26 zeigt die Anteile der absorbierten Histidinpeptide in der $Zn(Ri)_2$-Schicht. Die His_6-Peptide stehen zu 58 % in Kontakt zu der $Zn(Ri)_2$-Schicht. Dieser Anteil ist größer als derjenige in den Simulationen der einzelnen Histidinmoleküle. Das gleiche Ergebnis ist für die His_8-Peptide sichtbar. Die Hälfte aller eingesetzten Peptide ist in der $Zn(Ri)_2$-Schicht an die Zinkacetatfragmente gebunden. Des Weiteren ist ersichtlich, dass die Peptide aufgrund ihrer Hydrophobizität aggregieren. Der gebundene Anteil ist 8 % geringer als bei den His_6-Peptiden, welches auf die Aggregation der Peptide zurückzuführen ist. 55 % aller His_{10}-Peptide stehen in Kontakt mit dem $Zn(Ri)_2$. Insgesamt folgt hieraus, dass alle Histidinpeptide eine höhere Affinität zu der $Zn(Ri)_2$-Schicht haben als die einzelnen Histidinmoleküle selbst.

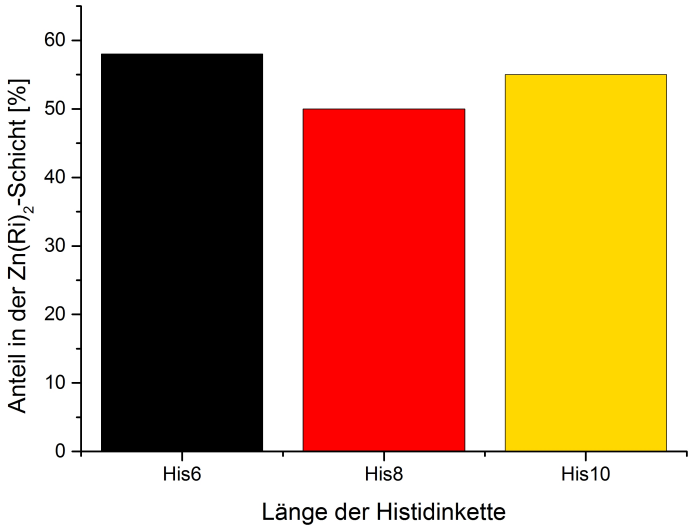


Abbildung 7.26: *Anteil der Moleküle in der $Zn(Ri)_2$-Schicht in Abhängigkeit von der Histidinkettenlänge.*

7.3.3 Interaktion von Insulin mit His-Tag und als Wildtyp mit einer $Zn(Ri)_2$-Schicht

Methode

Für die Untersuchung, ob Zinkricinoleat als neues Reagenz für die Aufreinigung von Peptiden und Proteinen geeignet ist, soll die Interaktion von Insulin als Wildtyp und in Konkurrenz mit einem His-Tag-Insulin an einer $Zn(Ri)_2$-Schicht untersucht werden. Dafür wird eine Simulationsbox mit den Maßen 28·28·21 nm^3 konstruiert. Die Box beinhaltet 10^5 Wassermoleküle und 1000 $Zn(Ri)_2$ Moleküle, welche am Boden der Box in einer xy-Schicht orientiert sind. In drei Simulationsboxen werden je ein Insulin des Wildtyps (PDB ID: 4INS[200]) und je ein Insulinmolekül mit einem His_6-Tag als auch mit einem His_8-Tag 4 nm über einer $Zn(Ri)_2$-Schicht platziert. Abbildung 7.27 zeigt die Startgeometrie, wobei die Imidazolfragmente des Histidins gelb, alle restlichen Fragmente des Insulins blau gefärbt sind. Die Stabilisierungskonstante der Proteinrückgratpotentiale ist auf $k_{BB} = 2.0$ gesetzt. Der pH-Wert der Simulation beträgt 7.4.

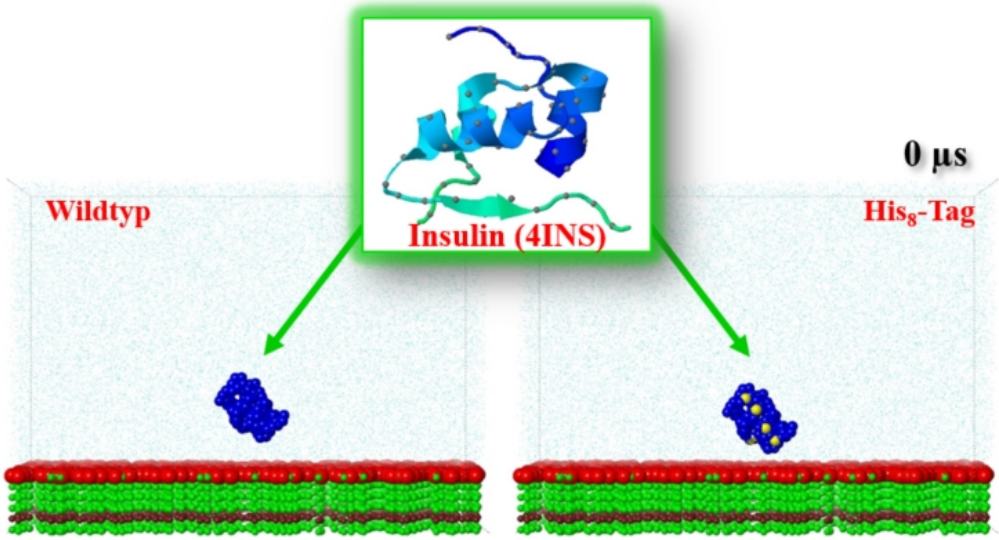


Abbildung 7.27: *Startgeometrie der Simulationsbox mit je einem Insulinmolekül.*

Eine weitere Simulationsbox wird erstellt, welche jeweils fünf Insulinmoleküle

des Wildtyps und fünf Moleküle mit His_8-Tag beinhaltet. Die Insulinmoleküle werden 4 nm über der $Zn(Ri)_2$-Schicht platziert, um den Interaktionsprozess von Zinkricinoleat mit dem Imidazolfragment in einer gemeinsamen Simulation studieren zu können. Abbildung 7.28 zeigt die Startgeometrie des Versuches.

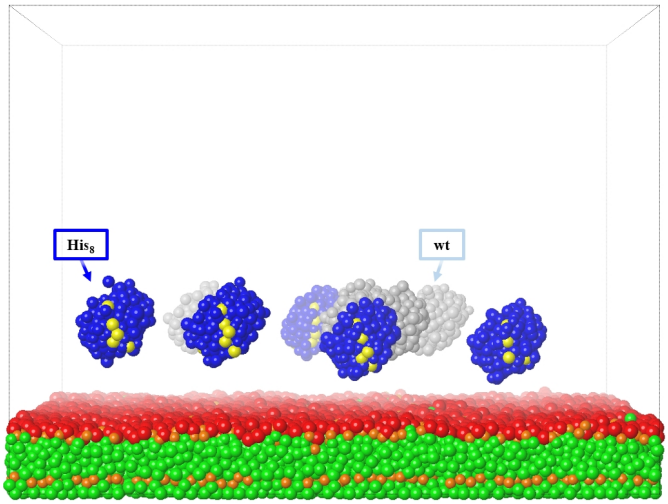


Abbildung 7.28: *Startgeometrie der Simulationsbox mit je fünf Insulinmolekülen des Wildtyps und fünf des His_8-Tag Insulinmoleküls.*

Ergebnis

Die vorangegangenen MFD-Simulationen konnten beweisen, dass Histidin und noch spezifischer Histidinpeptide eine erhöhte Affinität zu Zinkricinoleat aufweisen. An dieser Stelle soll gezeigt werden, dass die Interaktion mit Zinkricinoleat über die stickstoffhaltige Imidazolgruppe auch innerhalb von Proteinen (hier Insulin, PDB ID: 4INS[200]) in Form eines His-Tags stattfindet. In Abbildung 7.29 ist das Ergebnis der Simulationen mit der einzelnen Molekülspezies nach einer Simulationsdauer von 4.4 μs abgebildet. Eine permanente Assoziation des Wildtyp-insulins erfolgt nicht. Die Insulinmoleküle wechselwirken in beiden Fällen (His_6-Tag und His_8-Tag) sofort nach Simulationsbeginn mit der $Zn(Ri)_2$-Oberfläche. Der Kontakt besteht während der gesamten Simulation. Deutlich zu sehen ist in Abbildung 7.29 im vergrößertem Ausschnitt, dass die Interaktion über die gelb eingefärbten Imidazolfragmente vermittelt wird. Zu erkennen ist das Herauslösen eines Zinacetatfragments (rot) aus der $Zn(Ri)_2$-Schicht, wobei es von zwei Imida-

zolfragmenten umgeben ist. Des Weiteren war zu sehen, dass die Imidazolfragmente häufig die $Zn(Ri)_2$-Fragmente wechseln, sodass ein dynamischer Austausch von $Zn(Ri)_2$-Fragmenten stattfindet.

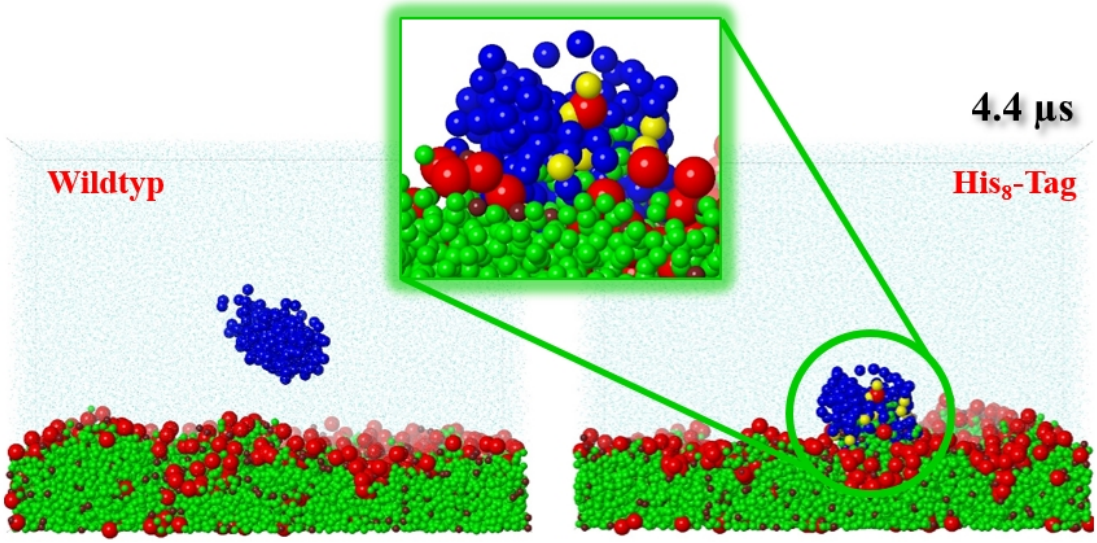


Abbildung 7.29: *Querschnitte der Simulationsbox mit je einem Insulinmolekül des Wildtyps und mit His_8-Tag.*

Im zweiten Versuch werden Moleküle des Wildtyps mit His_8-Tag-Molekülen gleichzeitig in einer Box simuliert. Die Wildtypinsulinmoleküle sind hierbei cyan gefärbt. Nach einer Simulationszeit von 2.2 μs werden 3 Insulinmoleküle mit His-Tag und 2 Moleküle des Wildtyps von der $Zn(Ri)_2$-Oberfläche angezogen. Am Ende der Simulation (4.4 μs) stehen alle His_8-Tag Moleküle in engem Kontakt mit der $Zn(Ri)_2$-Schicht. Ein Wildtypinsulin wird noch nicht vollständig an die Oberfläche adsorbiert.

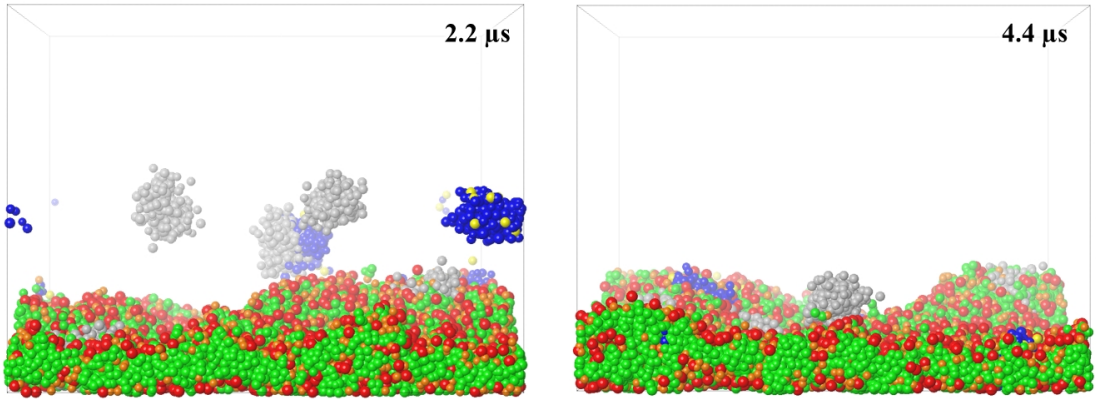


Abbildung 7.30: *Querschnitte der Simulationsbox mit je fünf Insulinmolekülen des Wildtyps und fünf des His_8-Tag Insulinmoleküls.*

Diese Ergebnisse zeigen, dass der vorgeschlagene Mechanismus der Interaktion des Zinkricinoleat über die stickstoffhaltigen funktionellen Gruppen, besonders derjenigen des Imidazolfragmentes der Aminosäure Histidin, vermittelt wird. Es konnte jedoch keine Bevorzugung der His-Tag Insulinmoleküle beim Absorptionsverhalten dargestellt werden.

7.4 Untersuchung der Membraninteraktion des Cyclotids Kalata B1

Der Prototyp aller Cyclotide, das Kalata B1, ist nur ein Mitglied einer großen Familie von Proteinen, welche ungefähr 50000 Proteine umfasst [89]. Die Zytotoxizität der Cyclotide ist auf ihre membranzerstörenden Eigenschaften zurückzuführen. Die hohe Affinität zu Lipidmembranen entsteht durch den einzigartigen Aufbau dieser Pflanzendefensine. Die Darstellung der Membranbindung sowie Mechanismen, die zu einer Membranzerstörung führen, sollen nachfolgend mittels MFD-Simulationen herausgearbeitet werden.

Methode

Für die Untersuchung der Wechselwirkung von Kalata B1 (PDB ID: 1NB1[201]) mit biologischen Membranen werden Simulationsboxen der Größe 32·32·32 nm^3 erstellt. Die Boxen beinhalten $2 \cdot 10^5$ Wassermoleküle und 1600 Lipide einer Plasmamembran. Die Lipide der Plasmamembran werden in einer xy-Schicht im unteren Drittel der Simulationsbox platziert. Die Kalata B1 Moleküle werden ebenfalls zufällig in einer xy-Schicht angeordnet, welche 5 nm über der Membranschicht positioniert ist. Für die Simulation wird die Einstellung *Sequence loop* aktiviert, um das geschlossene ringförmige Proteinrückgrat abzubilden. Folgende Kalata B1 Konzentrationen werden untersucht: 24, 48, 72, 96, 120, 144, 168, 192, 216, 240, 264, 288, 312, 336 und 350 Moleküle. Abbildung 7.31 skizziert die Startgeometrie der Simulationsboxen.

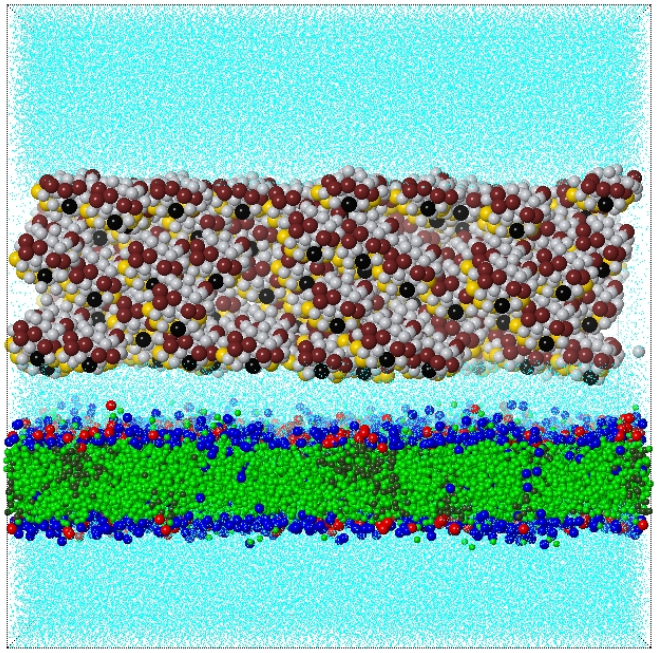


Abbildung 7.31: *Startgeometrie der Kalata B1 Simulationen.*

Damit die Ausrichtung der Kalata B1 Moleküle bzw. die Orientierung der funktionellen Gruppen des Proteins verfolgt werden können, sind die einzelnen Aminosäuren speziell eingefärbt. Die Aminosäuren der hydrophilen Aminosäurekette in Schlaufe 2 und 3 sind braun, die Aminosäuren des *hydrophobic Patch* (HP)

gelb und die beiden geladenen Aminosäuren mit Kontakt zu den geladenen Kopfgruppen der Phospholipide sind schwarz gefärbt. Für eine bessere Übersicht werden nur die Fragmente des Proteinrückgrates gezeichnet. Die Fragmente der Aminosäurenseitenketten sind ausgeblendet. Abbildung 7.32 stellt das Farbschema der Kalata B1 Moleküle dar.

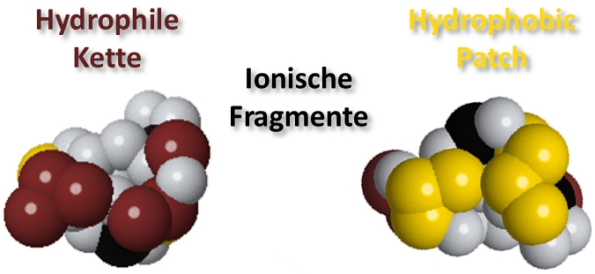


Abbildung 7.32: *Farbschema der Kalata B1 Moleküle innerhalb der Simulation (links: Ansicht von unten; Rechts: Ansicht von oben).*

Vor dem Start der jeweiligen Simulation werden 300 Minimierungsschritte durchgeführt. Der pH-Wert für die Ionisierung der Kalata B1 Moleküle beträgt 7.4. Die Kraftkonstante der Stabilisierungspotentiale beträgt $k_{BB} = 0.1$. Die Simulationszeit umfasst 35.5 μs. Dies entspricht $4{\cdot}10^5$ Dynamikschritten. Die Parameter der elektrostatischen Kraft sind in Tabelle 7.6 aufgeführt.

Tabelle 7.6: *Simulationsparameter der elektrostatischen Kraft.*

Parameter	Wert [MFD]
F^E_{max}	100
$r_{damping}$	5
$c_{damping}$	1
r_{cutoff}	5

Ergebnis

Betrachtet man die Seitenansicht einer Plasmamembran mit 120 Kalata B1 Molekülen in Abbildung 7.33, so sieht man, dass die Fragmente des HP (gelb) in Richtung der hydrophoben Schicht der Membran orientiert sind. Die hydrophile Kette (braun) ist in Richtung der wässrigen Phase ausgerichtet. Die geladenen Fragmente (schwarz) halten sich überwiegend auf Höhe der geladenen Kopfgruppen der Phospholipide auf. Die Draufsicht sowie die Ansicht von unten geben das selbe Bild wieder. In der Draufsicht ist zu erkennen, dass der überwiegende Anteil der hydrophilen Kette in Richtung der Wasserschicht zeigt. Die Ansicht von unten veranschaulicht den Kontakt des HP mit der hydrophoben Schicht der Membran, welche zur besseren Darstellung ausgeblendet ist. Dieses Bild entspricht dem theoretischem Bindungsmodell von Kalata B1 Molekülen an eine Membran, wie auch in der Literatur beschrieben[89–94,162].

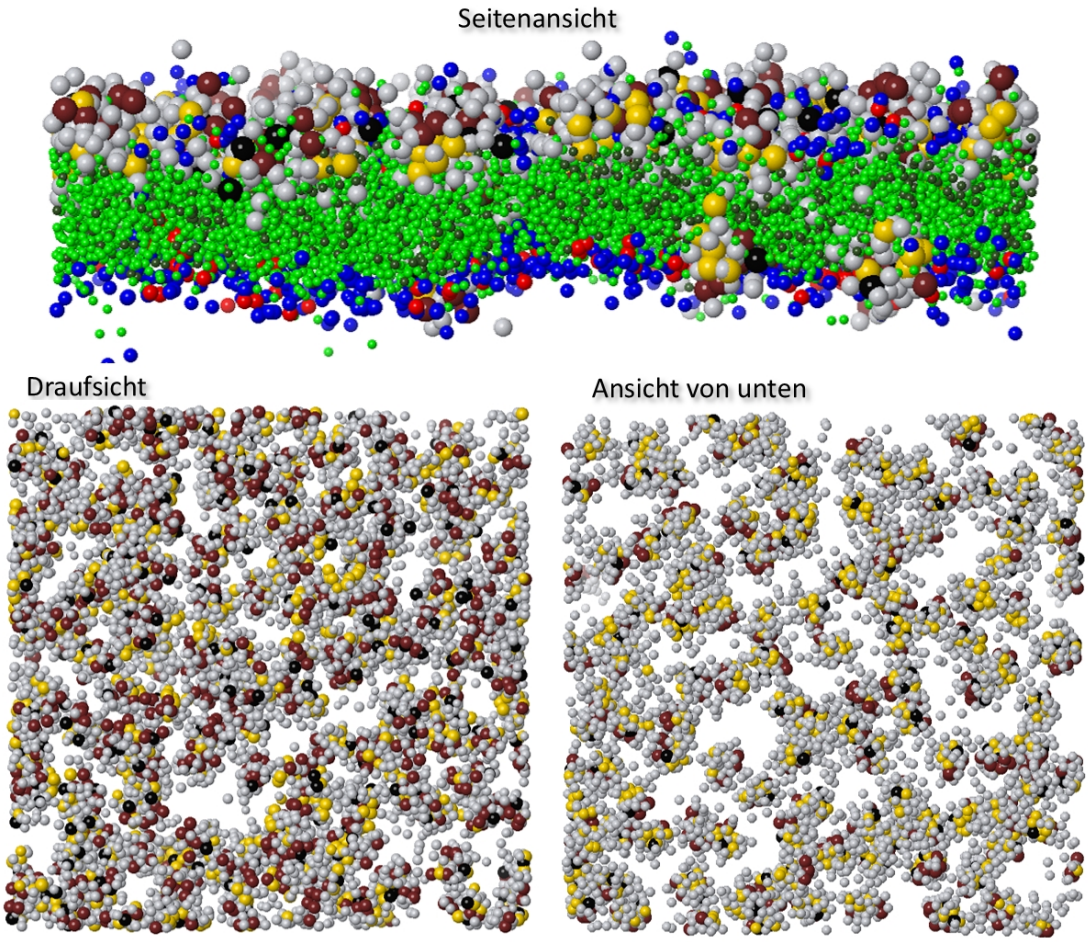


Abbildung 7.33: *Orientierung der Kalata B1 Moleküle bei Kontakt mit einer Plasmamembran. Die Membran ist in den unteren Abbildungen ausgeblendet.*

Um die Membrankrümmung in Abhängigkeit von der Konzentration an Kalata B1 zu bestimmen, wird der Abstand der geladenen Kopfgruppenfragmente mit 5% der niedrigsten (negative Membrankrümmung) bzw. der höchsten (positive Membrankrümmung) Position in z-Richtung mit der durchschnittlichen Position aller geladenen Kopfgruppenfragmente in z-Richtung verglichen. In Abbildung 7.34 ist das Ergebnis dieser Untersuchung dargestellt. Die blaue Kurve 'Min' bezieht sich auf die minimalen Punkte der negativen Membrankrümmung, die schwarze Kurve 'Max' auf die positive Membrankrümmung in Abhängigkeit der Kalata B1 Konzentration. Der Nullpunkt der linken y-Achse ist der Schwerpunkt der Membran. Auf der rechten Achse ist die Amplitude (rote Kurve) abgebildet, d.h. die Differenz zwischen der negativen und positiven Membrankrümmung. Bei einer

Konzentration von 24 Kalata B1 Molekülen beträgt die Amplitude 66.4 Å. Diese steigt langsam bis zu einer Konzentration von 192 Kalata B1 Molekülen auf einen Wert von 72.6 Å an. Ab einer Konzentration von mehr als 200 Kalata B1 Molekülen in der Simulationsbox vergrößert sich die Amplitude sprunghaft von 81.0 Å für 216 Moleküle bis auf einen Wert von 100.0 Å für 350 Moleküle.

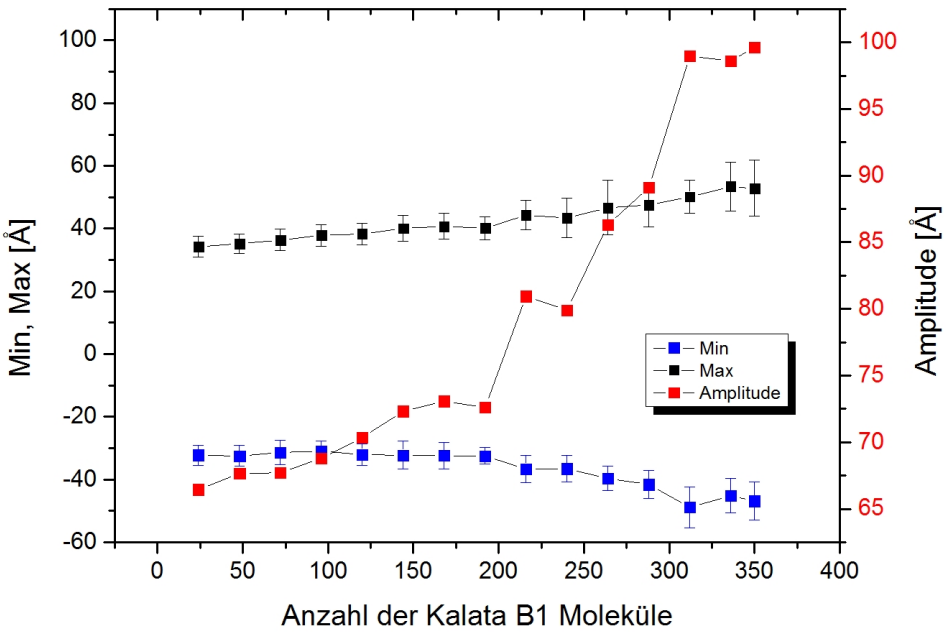


Abbildung 7.34: *Darstellung der negativen Membrankrümmung 'Min', der positiven Membrankrümmung 'Max' und der Amplitude in Abhängigkeit der Kalata B1 Konzentration.*

Der sprunghafte Anstieg der Membrankrümmmung ist ebenso anhand von Querschnitten durch die Simulationsbox ersichtlich. In Abbildung 7.35 sind Querschnitte der Simulationsboxen für 144, 264 und 350 Kalata B1 Moleküle dargestellt. Für jede Konzentration sind Aufnahmen zum Zeitpunkt 0.09 μs, 3.7 μs und 36.8 μs abgebildet. Nach 0.09 μs zeigt die Membran für jede Kalata B1 Konzentration das gleiche Ergebnis. Die Membranen besitzen eine glatte Struktur, welche keine Wellen oder andere geometrische Verzerrungen aufweist. Zum Zeitpunkt 3.7 μs stellt sich ein anderes Bild dar. Die Membran bei einer Konzentration von 144 Molekülen hat sich in ihrer Form nicht wesentlich verändert. Ab einer Konzentrationen von 264 und 350 Molekülen ist deutlich eine ausgeprägte Membrankrüm-

mung zu erkennen. Außerdem treten bei einer Konzentration von 350 Molekülen zu diesem Zeitpunkt extrahierte Lipide auf. Am Ende der Simulation nach 36.8 μs ist eine leichte Krümmung der Membran für 144 Kalata B1 Moleküle erkennbar. Die Membran der Konzentration von 264 Molekülen zeigt einen deutlichen wellenartigen Charakter. Die Betrachtung der Dynamiksimulation verdeutlicht die Fluktuationen in der Struktur der Membran. Dieser Eindruck verstärkt sich noch bei einer Konzentration von 350 Kalata B1 Molekülen. Auch hier ist eine Wellenstruktur in der Membran sichtbar. Zusätzlich zur Membrankrümmung im Querschnitt erkennt man eine Verdrehung der Membran in Richtung der Tiefe der Simulationsbox. Die Zunahme der Amplitude bzw. die Verstärkung der wellenförmigen Struktur der Membran ist von Nawae et al.[96] entsprechend mit *coarse-grained* MD-Simulationen abgebildet worden.

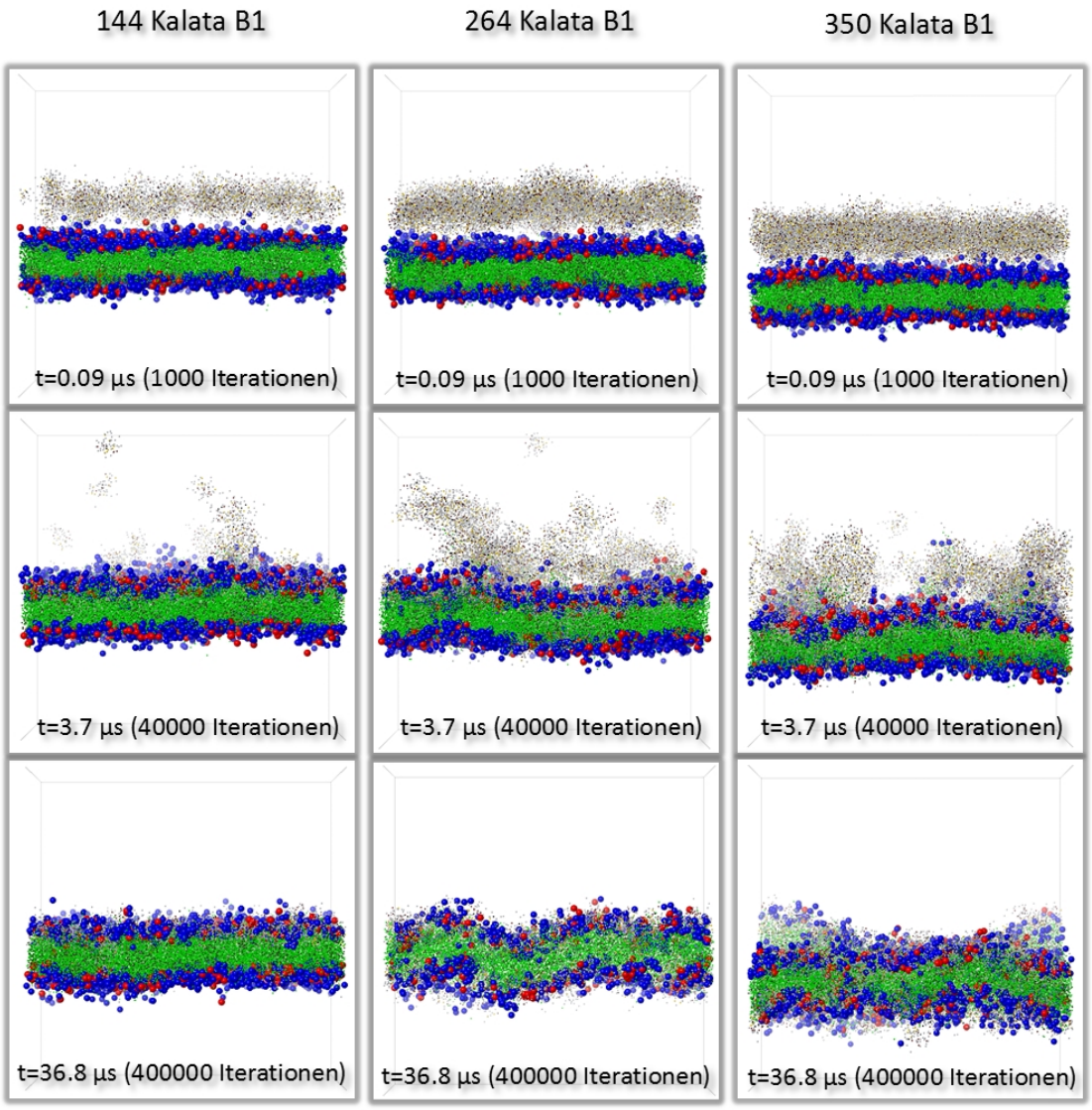


Abbildung 7.35: *Membranquerschnitte von verschiedenen Konzentrationen an Kalata B1 Molekülen (144, 264 und 350). Die Querschnitte werden zum Zeitpunkt* 0.09 *µs, 3.7 µs und 36.8 µs aufgenommen.*

Neben der Belastung der Membran als Folge der erhöhten Krümmung durch den Einfluss von Kalata B1 tritt ein weiteres Phänomen in Erscheinung. Ab einer Konzentration von mehr als 300 Kalata B1 Molekülen erfolgt die Extraktion von Lipidmolekülen aus der Membran. Abbildung 7.36 zeigt zwei Querschnitte mit Lipidextraktion nach einer Simulationszeit von 3.7 μs und 36.8 μs. Die linke Abbildung stellt den Beginn der Extraktion dar. Erkennbar ist, wie Kalata B1 Moleküle die Lipide in Form einer Säule aus der Membran heraus zieht. Auf der rechten Seite der Abbildung sieht man das Ende der Simulation. Große globuläre Einschlüsse an

Lipidmolekülen innerhalb von Kalata B1 haben sich über der Membran gebildet. Dies führt zu einer deutlichen Verringerung an Substanz innerhalb der ursprünglichen Membran.

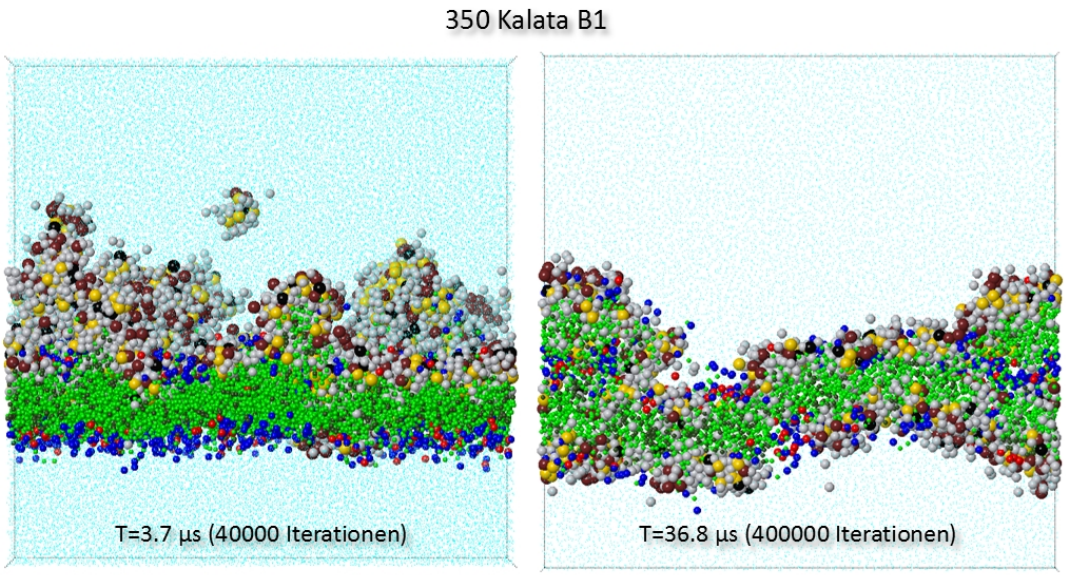


Abbildung 7.36: *Membranquerschnitte zur Betrachtung von Lipidextraktionen durch Kalata B1. Links nach 3.7 µs und rechts nach 36.8 µs Simulationszeit.*

Um eine Betrachtung der von Kalata B1 verursachten Löcher innerhalb der Membran zu ermöglichen, muss die Lipidanzahl um ca. 10 % auf 1450 Lipide verringert werden. Abbildung 7.37 stellt ein gebildetes Loch bei einer Konzentration von 216 Kalata B1 Molekülen dar. Das linke Bild zeigt die Membran mit Kalata B1 und rechts den gleichen Ausschnitt ohne das Protein. Im oberen Bild sieht man eine vergrößerte Ansicht des Loches, wobei die fehlenden Kalata B1-Moleküle als roter Kreis dargestellt sind. Die Bilder präsentieren die Simulationsboxen nach einer Zeit von 26.9 μs. Das Loch besitzt einen ungefähren Durchmesser von 8.0 nm, welches fast doppelt so groß ist, als die in der Literatur beschriebenen 4.1 - 4.7 nm[96]. Des Weiteren werden in der Literatur zwei Arten von Porenstrukturen postuliert: Die Fasslamellenpore sowie die Ringkernpore[89,92,95]. Betrachtet man die Orientierung und Lage der Kalata B1 Moleküle sowie der Phospholipide (hier speziell die Kopfgruppen), so erkennt man, dass hier eine Ringkernpore

ausgebildet ist. Innerhalb der Pore zeigen die *hydrophobic Patches* der Kalata B1 Moleküle in Richtung der Membran während die hydrophile Kette in Richtung des Poreninneren deutet. Die Kopfgruppen der Lipide sind zusätzlich in Richtung des Porenzentrums orientiert (siehe Abbildung 7.37 roter Kreis).

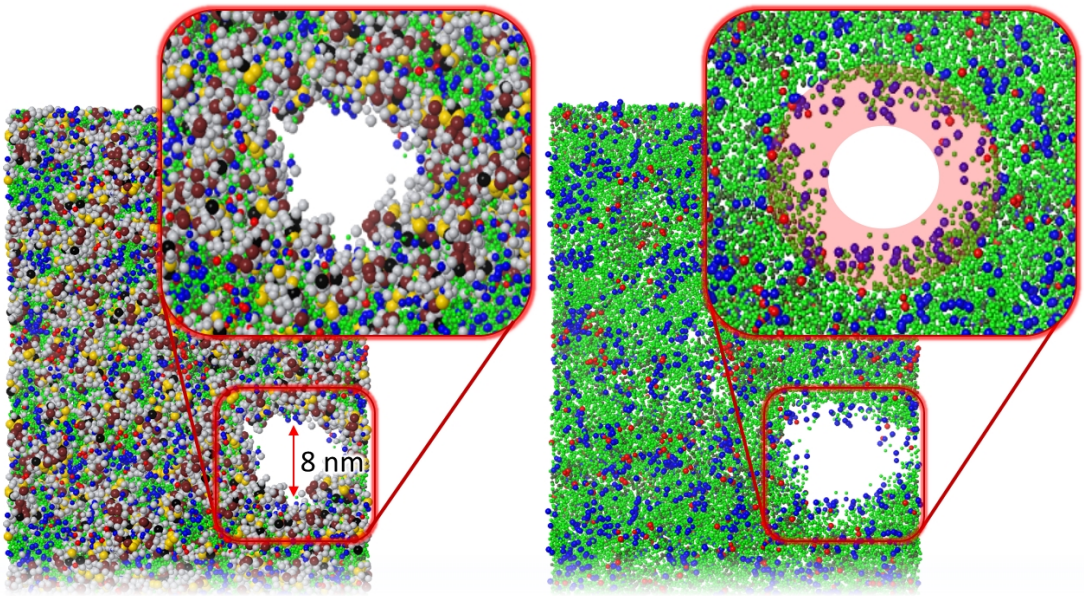


Abbildung 7.37: *Darstellung einer durch Kalata B1 hervorgerufenen Pore in einer Plasmamembran. Oben ist die Pore für beide Darstellungen vergrößert. Auf der rechten Seite ist das Protein ausgeblendet, oben die Position der ausgeblendeten Kalata B1 Moleküle rot hervorgehoben.*

Die hier durchgeführten MFD-Simulationen bilden die in der Literatur beschriebenen Abläufe der Membranbindung sowie Membranzerstörung exakt ab. Es konnte gezeigt werden wie Kalata B1 den Kontakt mit der Membran durchführt, die Krümmung der Membran bis hin zur Lipidextraktion beeinflusst, und Ringkernporen in der Membran ausbildet.

7.5 Diskussion

Die molekulare Fragmentdynamik approximiert anisotrope molekulare Wechselwirkungen in Form eines isotropen Repulsionsparameters. Die Struktur von Prote-

inen wird mittels anisotropen Wechselwirkungen wie den Wasserstoffbrückenbindungen stabilisiert. Aus diesem Grund werden im Rahmen dieser Arbeit harmonische Stabilisierungspotentiale für das Proteinrückgrat eingeführt. Anhand des Beispiels Calmodulin kann verdeutlicht werden, dass eine erfolgreiche Stabilisierung der Proteinkonformation mit Hilfe der Stabilisierungspotentiale möglich ist. Zusätzlich kann die Flexibilität der Geometrie des Proteins über die Kraftkonstante der harmonischen Potentiale fein abgestimmt werden. Somit können Proteine mit vollständiger Strukturkonservierung, aber auch mit einer hohen Flexibilität untersucht werden. Eine Restriktion der Reichweite der Potentiale ermöglicht die Konservierung der Sekundärstruktur, ohne die gesamte Struktur in einen Potentialkäfig zu sperren. Dabei konnte bewiesen werden, dass die langreichweitigen Potentiale die wichtigste Komponente für die Stabilisierung der Geometrie darstellen.

Die Untersuchung der Stabilität der Quartärstruktur des Proteinkomplexes Hämoglobin legt dar, dass die molekulare Fragmentdynamik die globuläre Struktur des Moleküls aufrecht erhält. Nur kurzzeitig bildet sich eine planare Konformation in Form eines metastabilen Übergangszustandes aus. Die exakte Quartärstruktur der kristallographischen Daten wird während der Simulation nicht eingehalten. Dies veranschaulicht die Auswertungen der jeweiligen Sondenpaare bezüglich der Distanz und Winkelveränderungen untereinander. Die Untereinheiten sind nicht starr verknüpft, sondern frei beweglich. Dennoch bleibt die ungefähre Orientierung der Hb α- und Hb β-Untereinheiten zueinander erhalten. Die einzelnen Untereinheiten stehen in ständigem Kontakt miteinander. Zusammenfassend lässt sich daraus schließen, dass das Fehlen von gerichteten Wechselwirkungen wie den Wasserstoffbrückenbindungen die exakte Struktur nur schwer aufrecht erhalten lässt. Die molekulare Fragmentdynamik beschreibt die Wechselwirkungen gemittelt und in korrekter Weise, da die vorgegebene Orientierung der Untereinheiten zueinander ersichtlich bleibt.

Die erste konkrete Anwendung der hier vorgestellten Methode zur Beschreibung von Peptiden und Proteinen ist eine Studie des Moleküls Zinkricinoleat ($Zn(Ri)_2$). Eine Eigenschaft des Zinkricinoleat ist es, mit Geruchsstoffen über deren schwefel-

und stickstoffhaltige funktionelle Gruppen zu wechselwirken[145]. Interessant ist, ob Zinkricinoleat die Fähigkeit besitzt, mit der Aminosäure Histidin, welche eine Imidazolgruppe enthält, zu interagieren.

Als erste wird eine MFD-Simulation durchgeführt, in welcher ein Histidinmolekül über einer $Zn(Ri)_2$-Schicht platziert wird. Während der gesamten Simulation findet keine Dissoziation des Histidins von der $Zn(Ri)_2$-Oberfläche statt. Die Wechselwirkung entsteht über das Zinkacetatfragment und dem Imidazolfragment des Histidins, wie in experimentellen und theoretischen Studien vorgeschlagen[145,202]. Um die Selektivität von Zinkricinoleat für Histidin zu testen, werden verschiedene Mischungen von Histidin und einer anderen Aminosäure über einer $Zn(Ri)_2$-Schicht platziert. Das Ergebnis zeigt, dass die Imidazolfragmente in Kontakt mit den Zinkacetatfragmenten der $Zn(Ri)_2$-Oberfläche treten und dort festgehalten werden. Im Gegensatz dazu treten die anderen Aminosäuren in Kontakt mit dem hydrophoben Anteil der $Zn(Ri)_2$-Schicht. Dies zeigt, dass die Interaktion des Imidazolfragmentes mit den Zinkacetatfragmenten imstande ist, die Histidinmoleküle an der Oberfläche des Zinkricinoleats zu halten. Dies belegt die dokumentierten Experimente, dass nukleophile Substanzen wie Ammonium, Amine und Schwefelwasserstoffe selektiv von Zinkricinoleat absorbiert werden[145,202].

Im zweiten Versuch mit Zinkricinoleat wird die Affinität von Polyhistidinpeptiden (kurz His_N, wobei N die Kettenlänge beschreibt) analysiert. Diese Untersuchung dient der Vorstudie zur Ermittlung der Möglichkeiten zur Aufreinigung von His-Tag Proteinen mittels $Zn(Ri)_2$ wie in Patent DE102011101880 A1 beschrieben. Der Boden der Box besteht aus einer $Zn(Ri)_2$-Schicht. Über die $Zn(Ri)_2$-Schicht werden 100 Polyhistidinmoleküle mit verschiedenen Kettenlängen (His_6, His_8 und His_{10}) platziert. Nach einer Simulationszeit von 4.4 μs sind 58 % der His_6-, 50 % der His_8- und 55 % der His_{10}-Proteine gebunden. Die Affinität der Polyhistidinpeptide ist somit höher als die der einzelnen Histidinmoleküle. Auffällig ist die Aggregation der Peptide untereinander, wodurch aufgrund der Größe der Aggregate die Diffusion zur $Zn(Ri)_2$-Schicht verlangsamt wird. Die Wechselwirkung wird über das Imidazolfragment der Histidinmoleküle mit dem Zinkacetatfragment vermittelt, wie experimentell belegt[145,202].

Das Patent DE102011101880 A1 beschreibt, dass Zinkricinoleat selektiv His-Tag-Proteine binden kann. Somit könnte Zinkricinoleat als neues Auftrennungsmedium fungieren und als Alternative zur Metallchelatchromatographie (IMAC) dienen. Um die Aufreinigungsfähigkeiten für Proteine zu demonstrieren, werden Simulationen mit Insulin des Wildtyps und Insulin mit einem His-Tag durchgeführt, wobei der His-Tag aus sechs und acht Histidinmolekülen (His_6 und His_8) besteht. Deutlich zu erkennen bei diesem Versuch ist die vermehrte Wechselwirkung der Imidazolfragmente mit den Zinkacetatkopfgruppen der $Zn(Ri)_2$-Schicht. Um zu untersuchen, ob der Aufreinigungsprozess selektiv ist, werden Insulinmoleküle des Wildtyps mit His_8-Tag Molekülen gemeinsam in einer Box platziert. Bis auf ein einzelnes Wildtypmolekül werden alle Moleküle von der $Zn(Ri)_2$-Schicht absorbiert. Aus alledem folgt, dass Zinkricinoleat zwar eine deutlich erhöhte Affinität zum Imidazolfragment des Histidins aufweist, jedoch keine besondere Bevorzugung gegenüber anderen Aminosäuren vorliegt. Die eigentliche Absorption wird über eine Komplexgleichgewichtsreaktion vermittelt, welche mit der molekularen Fragmentdynamik nicht abgebildet werden kann. Jedoch zeigt sich tendenziell eine erhöhte Affinität des Histidins zum Zinkricinoleat, speziell zur Zinkacetatkopfgruppe. Des Weiteren ermöglichen die weichen Fragmentpotentiale ein Durchdringen der einzelnen Moleküle, sodass Proteine ohne His-Tag mit der hydrophoben Schicht innerhalb der $Zn(Ri)_2$-Schicht in Kontakt treten und so in diese eindringen. Dies wird bei der Verwendung von harten Potentialen weitestgehend ausgeschlossen.

Cyclotide sind aufgrund ihrer Vielfalt an biologischen Aktivitäten interessante Vorschungsobjekte. Als Pflanzendefensine[92] entsteht ihre Zytotoxizität durch ihre membranzerstörenden Eigenschaften. Eine genauere Untersuchung des Wirkungsmechanismus mittels hier durchgeführter MFD-Simulationen hat gezeigt, dass der vorgeschlagene Ablauf des Membrankontaktes des Cyclotids Kalata B1 über den *hydrophobic Patch* (HP) und den beiden geladenen Aminosäureresten (G_6 in Schlaufe 1 und N_{29} in Schlaufe 6) vermittelt wird[89–94]. Der HP steht in Kontakt mit der hydrophoben Schicht der Membran, die beiden ionischen Aminosäurereste halten sich exakt in der Schicht der geladenen Kopfgruppen auf.

Der wichtigste Mechanismus, welcher zur Perforation der Zellmembran führt, ist die Ausübung von Stress in Form von Membrankrümmung. Dies führt zur erleichterten Bildung von Poren durch die Cyclotide oder zum Zerreißen der Lipiddoppelschicht[91]. Die Stärke der Membrankrümmung ist abhängig von der Konzentration der Cyclotide[96]. Für eine Konzentration von weniger als 200 Kalata B1-Molekülen auf 1600 Phospholipide ergibt sich nur eine geringe Steigerung der Amplitude (Differenz der negativen und positiven Krümmung der Membran) der Membrankrümmung. Die Erhöhung der Amplitude zwischen 24 und 196 Kalata B1 Molekülen liegt bei 8.5 %. Eine Konzentration von 216 zeigt einen Amplitudenzuwachs von 18.0 %. 312 Kalata B1 Moleküle kennzeichnen sich durch einen 32.9 % größeren Wert der Amplitude, welches einer ungefähren Vervierfachung gegenüber dem Wert von 196 Kalata B1 Molekülen entspricht. Die im Rahmen dieser Arbeit durchgeführten MFD-Simulationen konnten zeigen, dass das Ausmaß der Membrankrümmung direkt abhängig von der Kalata B1 Konzentration ist.

Für höhere Konzentrationen ist in der Literatur vermerkt, dass Lipidmoleküle direkt aus der Doppelschicht extrahiert werden[91,96]. Dieses Phänomen kann für Konzentrationen größer als 300 Kalata B1 Moleküle beobachtet werden. Zu Beginn der Simulation (3.7 μs) werden säulenartige Lipidanhäufungen durch die Kalata B1 Schicht herausgezogen. Zum Ende hin (36.8 μs) sind diese Lipidanhäufungen zu globulären Einschlüssen kleiner 150 Å herangewachsen. Dies hat eine deutliche Verringerung der Membrandicke zur Folge.

In der Literatur sind zwei Arten von Poren, die Kalata B1 ausbildet, erwähnt: Die Fasslamellenpore sowie die Ringkernpore[89,92,95]. In einer hier durchgeführten Simulation konnte bei einer um 10 % reduzierten Anzahl an Phospholipiden gezeigt werden, dass die ausgebildete Pore einer Ringkernporenstruktur entspricht. Die Größe der Pore ist mit 8 nm größer als die in der Literatur beschriebenen 4.1 - 4.7 nm[96]. Fasst man die gesammelten Ergebnisse zusammen, so konnte bewiesen werden, dass die molekulare Fragmentdynamik fähig ist, die komplexen Mechanismen der Membraninteraktion von Kalata B1 zu veranschaulichen.

8 Zusammenfassung

Die dissipative Partikeldynamik (DPD) ist eine etablierte mesoskopische Simulationsmethode für komplexe Fluidsysteme, deren Partikel geeignete Fluidelemente repräsentieren. Dabei werden die konservativen Kräfte zwischen Partikelpaaren durch jeweilige isotrope Repulsionsparameter quantifiziert. Die molekulare Fragmentdynamik (MFD) ist eine DPD-Variante, die kleine Moleküle als DPD-Partikel definiert. Größere Moleküle werden dann aus Partikel-„Fragmenten“ zusammengesetzt, die über harmonische Federn verbunden sind. Die MFD-Methode wurde bereits erfolgreich für die Untersuchung von amphiphilen Polymeren, Mikroemulsionen und nichtionischen Tensiden eingesetzt. Ziel der vorliegenden Arbeit ist es, eine Erweiterung und Validierung des MFD-Ansatzes für die Simulation biomolekularer Systeme mit Membranen, Peptiden oder Proteinen zu realisieren.

Dazu wird zunächst ein Satz von MFD-Fragmenten definiert, der eine geeignete Fragmentierung der in den Biomembranen auftretenden Phospholipide sowie der die Peptide/Proteine konstituierenden Aminosäuren ermöglicht. Für die Beschreibung der jeweiligen fragmentbasierten Molekültopologie wird eine neue einzeilige Textnotation (*fragment* SMILES oder kurz *f*SMILES) eingeführt, die sich an der bekannten SMILES-Kodierung orientiert. Die DPD-Repulsionsparameter für alle auftretenden Fragmentpaare werden mit Molekulardynamiksimulationen bestimmt und mit einer neu entwickelten Verwaltungsapplikation dokumentiert. Ein Ad-hoc-Konzept für die Berücksichtigung elektrostatischer Ladungen und Kräfte in MFD-Simulationen wird vorgeschlagen, getestet und eingesetzt. Speziell für Peptide und Proteine werden harmonische Federkräfte zwischen den Rückrad-Fragmenten implementiert, die eine Stabilisierung ihrer räumlichen 3D-Struktur mit variablen Flexibilitätsgraden ermöglichen, da die isotropen DPD-Partikel-

wechselwirkungen die tatsächlichen räumlich-gerichteten atomaren Wechselwirkungen prinzipiell nicht adäquat widerspiegeln. Für die praktische Durchführung der biomolekularen Simulationen wird die MFD-Simulationssoftware MFD-FormulaOne mit umfangreichen neu entwickelten Funktionsbibliotheken und (Peptid-/Protein-)Editoren erweitert. Dabei wird nach Möglichkeit auf bestehende Open-Source-Bibliotheken (CDK, Jmol, BioJava) aufgesetzt. Das erweiterte Simulationssystem ermöglicht schließlich eine vollständig automatisierte Verarbeitung der in „Protein-Data-Bank“-Einträgen enthaltenen topologischen und räumlichen Peptid-/Protein-Information.

Zur ersten Validierung der erweiterten Software werden zunächst die Polyoxyalkylethertenside C_6E_6, $C_{10}E_6$, $C_{12}E_6$ und $C_{16}E_6$ an einer Wasser-Luft-Grenzfläche simuliert. Die erhaltenen nanoskaligen Strukturen im thermodynamischen Gleichgewicht, die berechneten Oberflächenspannungen sowie die Ergebnisse für die Mizellenbildung weisen eine gute Übereinstimmung mit experimentellen Ergebnissen auf.

Für die Untersuchung von Biomembranen werden geeignete Fragmentierungsschemata für alle auftretenden Phospholipide erarbeitet. Die Simulation von 1,2-Dimyristoyl-sn-glycero-3-phosphatidylcholin-(DMPC)-, Endoplasmatischen-Retikulum-(ER)-, Mitochondrium- und Plasma-Modellmembranen führt zu stabilen Doppelschichten mit Geometrien, die experimentellen Untersuchungen gut entsprechen, wobei auch die bekannten Lipid-Fluktuationen zwischen den beiden Schichten der Doppelmembranen auftreten. Mizellenbildung sowie spontane Vesikelbildung können für alle untersuchten Membrantypen simuliert werden. Die Fusion eines DMPC-Vesikels mit einer DMPC-Membran wird zudem erfolgreich dargestellt.

Die flexible räumliche Stabilisierung des Proteinrückgrads mit harmonischen Federkräften wird für das hantelförmige Calmodulin eingehend untersucht: Dazu werden Simulationsserien mit unterschiedlichen Federkräften durchgeführt und die dazu gehörenden räumlichen Proteingeometrien ausgewertet. Es zeigt sich, dass die langreichweitigen Federkräfte in erster Linie für die Aufrechterhaltung der 3D-Struktur entscheidend sind. Die Stabilität einer Protein-Quartärstruktur wird

mittels einer MFD-Langzeitsimulation für die Aggregation der Untereinheiten des Hämoglobin-Tetramers gezeigt. Dies belegt, dass die MFD-Methode die gemittelte Wechselwirkung zwischen Proteinoberflächen adäquat beschreiben kann.

Für das Studium der Wechselwirkung von Peptiden/Proteinen und molekularen Oberflächen werden zwei Modellsysteme untersucht. Zum einen wird die bekannte destruktive Wirkung des Pflanzendefensins Kalata B1 auf eine biologische Plasmamembran erfolgreich simuliert und die Struktur auftretender Membranporen ermittelt. Zum anderen kann die Wechselwirkung von histidinhaltigen Peptiden und Proteinen mit einer technisch relevanten Zinkrizinoleatschicht die postulierte Affinität des Zinkrizinoleats zu stickstoffhaltigen Molekülen aufzeigen.

Zusammenfassend kann die in dieser Arbeit erweiterte MFD-Methode als neue Simulationstechnik für biomolekulare Systeme gewertet werden, die als mesoskopisch-grobkörniges Verfahren für geeignete Fragestellungen einen um Größenordnungen verringerten Rechen- und Zeitaufwand gegenüber der atomistischen Molekulardynamik ermöglicht.

9 Summary

Dissipative Particle Dynamics (DPD) is a well-established mesoscopic simulation method for complex fluid systems whose particles are represented by adequate fluid elements. The conservative forces between particle pairs are quantified by particular isotropic repulsion parameters. Molecular Fragment Dynamics (MFD) is a DPD variant that utilizes small molecules as its particles where larger molecules are assembled with these particle 'fragments' connected by harmonic springs. The MFD technique has already been successfully applied to study amphiphilic polymers, microemulsions and non-ionic surfactants. This thesis aims at extending and validating the MFD approach for the simulation of biomolecular systems containing membranes, peptides and proteins.

At first a set of MFD fragments is defined that allows an adequate fragmentation of all membrane relevant phospholipids as well as the amino acids that set up peptides and proteins. For a description of a particular fragment-based molecule topology a new single-line text notation (denoted *fragment* SMILES or *f*SMILES) is introduced that is similar to the well-known SMILES code. The DPD repulsion parameters for all fragment pairs are determined by molecular dynamics simulations and documented by a newly developed management application. An ad-hoc concept for the consideration of electrostatic charges and forces in MFD simulations is proposed, tested and applied. For peptides and proteins specific harmonic forces between backbone fragments are implemented that enable a stabilisation of their spatial 3D structures with variable degrees of flexibility since the isotropic DPD particle interactions cannot account for the true spatially directed atomic interactions. In order to practically realize biomolecular simulations the MFD simulation software MFD-FormulaOne is extended with new function libraries and

peptide/protein editors where already existing open-source libraries (CDK, Jmol, BioJava) are utilized according to possibility. The extended simulation system allows a completely automated processing of the topological and spatial peptide/-protein information contained in Protein Data Bank (PDB) files.

For an initial validation of the extended software the polyoxyalkylether surfactants C_6E_6, $C_{10}E_6$, $C_{12}E_6$ and $C_{16}E_6$ are simulated at a water-air interface. The resulting nanoscale structures at thermodynamic equilibrium, the calculated surface tensions and the observed micelle formations are in good agreement with experimental findings.

In order to study biomembranes adequate fragmentation schemata for all occurring phospholipids are developed. The simulation of 1,2-Dimyristoyl-sn-glycero-3-phosphatidylcholin (DMPC), Endoplasmic reticulum (ER), mitochondrion and plasma model membranes lead to stable bilayers with geometries that correspond to experimental results. The well-known lipid fluctuations between both membrane layers do appear. Micelle formation as well as spontaneous vesicle formation can be demonstrated for all studied membrane types. The fusion of a water-filled DMPC vesicle with a DMPC bilayer membrane is successfully simulated.

The flexible spatial stabilization of the protein backbones with harmonic forces is studied with the dumbbell-shaped Calmodulin: Simulation series with different forces are performed and the corresponding spatial protein geometries are analysed. It is shown that the long-range forces are responsible for the preservation of the 3D structure in the first place. The stability of a protein quaternary structure is demonstrated by a long-term MFD simulation of the aggregation of the subunits of the Hemoglobin tetramer. This shows that the MFD method may successfully describe the average interaction between protein surfaces.

To analyse the interaction between peptides/proteins and molecular surfaces two model systems are investigated: First the well-known destructive behaviour of Kalata B1 plant defensins with biological plasma membranes is successfully simulated and the structure of the occurring membrane pores is characterized. Secondly the interaction of Histidine containing peptides and proteins with a tech-

nical relevant zinc ricinoleate layer shows the postulated affinity of zinc ricinoleate to nitrogen containing molecules.

In summary it is shown that the extended MFD method may be regarded as a new simulation technique for biomolecular systems. Where it applies the mesoscopic coarse-grained MFD method may reduce the computational costs by orders of magnitude compared to atomistic molecular dynamics.

10 Literatur

[1] Macmillan, W. *Bulletin of the American Mathematical Society* **1913**, *11*, 349–355.

[2] Rahman, A. *Physical Review* **1964**, *136*, 405–411.

[3] Gibson, J.; Goland, A.; Milgram, M.; Vineyard, G. *Physical Review* **1960**, *120*, 1229–1253.

[4] Alder, B. J.; Wainwright, T. E. *The Journal of Chemical Physics* **1957**, *27*, 1208–1209.

[5] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *The Journal of Chemical Physics* **1953**, *21*, 1087.

[6] Moore, G. *Electronics* **1965**, *38*, 114–117.

[7] Facility, O. R. L. C. Titan. `http://www.olcf.ornl.gov/titan`, 2013.

[8] Friesner, R. A.; Dunietz, B. D. *Accounts of Chemical Research* **2001**, *34*, 351–358.

[9] Hansson, T.; Oostenbrink, C.; van Gunsteren, W. *Current Opinion in Structural Biology* **2002**, *12*, 190–196.

[10] Karplus, M.; McCammon, J. A. *Nature Structural Biology* **2002**, *9*, 646–652.

[11] Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *Journal of Computational Chemistry* **2000**, *21*, 1442–1457.

[12] Nobelprize.org: The Nobel Prize in Chemistry 1998. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/, Accessed: 23 Jan 2014.

[13] Nobelprize.org: The Nobel Prize in Chemistry 2013. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/, Accessed: 23 Jan 2014.

[14] Hsu, H.-P.; Grassberger, P. *Journal of Statistical Physics* **2011**, *144*, 597–637.

[15] Kolinski, A.; Skolnick, J. *Proteins* **1994**, *18*, 338–352.

[16] Skolnick, J.; Kolinski, A. *Science* **1990**, *250*, 1121–1125.

[17] Horejs, C.; Mitra, M. K.; Pum, D.; Sleytr, U. B.; Muthukumar, M. *The Journal of Chemical Physics* **2011**, *134*, 125103.

[18] Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. *Annual Review of Biophysics* **2012**, *41*, 429–52.

[19] Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341–346.

[20] Gunsteren, W.; Berendsen, H.; Geurtsen, R.; Zwinderman, H. *Annals of the New York Academy of Sciences* **1986**, *482*, 287–303.

[21] Shaw, D.; Dror, R. Millisecond-scale molecular dynamics simulations on Anton. High Performance Computing Networking, Storage and Analysis. 2009; pp 1–11.

[22] Durrant, J. D.; McCammon, J. A. *BMC Biology* **2011**, *9*, 71.

[23] van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708–28.

[24] Lin, H.; Truhlar, D. G. *Theoretical Chemistry Accounts* **2006**, *117*, 185–199.

[25] Groenhof, G. *Angewandte Chemie* **2013**, *125*, 12718–12720.

[26] Beierlein, F.; Lanig, H.; Schürer, G.; Horn, A. H. C.; Clark, T. *Molecular Physics* **2003**, *101*, 2469–2480.

[27] Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *Journal of Physics: Condensed Matter* **2004**, *16*, R481–R512.

[28] Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. *Journal of Structural Biology* **2007**, *157*, 593–605.

[29] Shi, Q.; Izvekov, S.; Voth, G. A. *The Journal of Physical Chemistry. B* **2006**, *110*, 15045–15048.

[30] Rudd, R.; Broughton, J. *Physical Review B* **1998**, *58*, R5893–R5896.

[31] de Pablo, J. J. *Annual Review of Physical Chemistry* **2011**, *62*, 555–574.

[32] Groot, R. D. *The Journal of Chemical Physics* **2003**, *118*, 11265–11277.

[33] Frenkel, D.; Smit, B. *...principles and practice of Monte Carlo and molecular ...*; Academic Press, Inc: San Diego, 2002; p 664.

[34] Hoogerbrugge, P.; Koelman, J. *EPL (Europhysics Letters)* **1992**, *19*, 155–160.

[35] Koelman, J.; Hoogerbrugge, P. *EPL (Europhysics Letters)* **1993**, *21*, 363–368.

[36] Pagonabarraga, I. *EPL (Europhysics Letters)* **1998**, *42*, 377–382.

[37] Pagonabarraga, I.; Frenkel, D. *The Journal of Chemical Physics* **2001**, *115*, 5015–5027.

[38] Flekkoy, E. G.; Coveney, P. V. *Physical Review Letters* **1999**, *83*, 1775–1778.

[39] Espanol, P.; Warren, P. *EPL (Europhysics Letters)* **1995**, *30*, 191–196.

[40] Groot, R. D.; Warren, P. B. *The Journal of Chemical Physics* **1997**, *107*, 4423–4435.

[41] Guo, X. D.; Zhang, L. J.; Wu, Z. M.; Qian, Y. *Macromolecules* **2010**, *43*, 7839–7844.

[42] Weininger, D.; Weininger, A.; Weininger, J. L. *Journal of Chemical Information and Computer Sciences* **1989**, *29*, 97–101.

[43] Weininger, D. *Journal of Chemical Information and Computer Sciences* **1988**, *28*, 31–36.

[44] Weininger, D. *Journal of Chemical Information and Computer Sciences* **1990**, *30*, 237–243.

[45] EMBL-EBI, UniProtKB/TrEMBL Protein Database Realease 2014_10 Statistics. `http://www.ebi.ac.uk/uniprot/TrEMBLstats/`, 2014.

[46] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Research* **2000**, *28*, 235–242.

[47] Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org.

[48] Truszkowski, A.; Epple, M.; Fiethen, A.; Zielesny, A.; Kuhn, H. *Journal of Colloid and Interface Science* **2013**, *410*, 140–145.

[49] Brini, E.; Algaer, E. a.; Ganguly, P.; Li, C.; Rodríguez-Ropero, F.; van der Vegt, N. F. a. *Soft Matter* **2013**, *9*, 2108–2119.

[50] Jacobson, S. H.; Gordon, D. J.; Nelson, G. V.; Balazs, A. *Advanced Materials* **1992**, *4*, 198–205.

[51] Fan, C. F.; Olafson, B. D.; Blanco, M.; Hsu, S. L. *Macromolecules* **1992**, *25*, 3667–3676.

[52] Blanco, M. *Journal of Computational Chemistry* **1991**, *12*, 237–247.

[53] Schulz, S.; Kuhn, H.; Schmid, G. *Materials Research Society* **2004**, *788*.

[54] MFD-FormulaOne, pre-release, version 1.0.0.85. MFD-FormulaOne is a product of CAM-D Technologies GmbH (www.molecular-dynamics.de) and GNWI - Gesellschaft fuer naturwissenschaftliche Informatik mbH (www.gnwi.de). http://www.mfd-formulaOne.de, 2014.

[55] Vishnyakov, A.; Lee, M.-T.; Neimark, A. V. *The Journal of Physical Chemistry Letters* **2013**, *4*, 797–802.

[56] Schulz, S. G.; Kuhn, H.; Schmid, G.; Mund, C.; Venzmer, J. *Colloid and Polymer Science* **2004**, *283*, 284–290.

[57] Ryjkina, E.; Kuhn, H.; Rehage, H.; Müller, F.; Peggau, J. *Angewandte Chemie International Edition* **2002**, *41*, 983–986.

[58] Ginzburg, V. V.; Chang, K.; Jog, P. K.; Argenton, A. B.; Rakesh, L. *The Journal of Physical Chemistry B* **2011**, *115*, 4654–4661.

[59] Li, Y.; Zhang, H.; Bao, M.; Wang, Z. *Computational Materials Science* **2012**, *63*, 154–162.

[60] Fan, H.; Striolo, A. *Physical Review E* **2012**, *86*, 051610.

[61] Li, Z. Q.; He, X. J.; Li, Y.; Ma, B. M.; Cao, X. L.; Song, X. W.; Cui, X. H. *Acta Chimica Sinica* **2007**, *65*, 2803–2808.

[62] Li, Y.; He, X.; Cao, X.; Zhao, G.; Tian, X.; Cui, X. *Journal of Colloid and Interface Science* **2007**, *307*, 215–220.

[63] Li, Y.; Zhang, P.; Dong, F.-L.; Cao, X.-L.; Song, X.-W.; Cui, X.-H. *Journal of Colloid and Interface Science* **2005**, *290*, 275–280.

[64] Gezae Daful, A.; Baulin, V. a.; Bonet Avalos, J.; Mackie, A. D. *The Journal of Physical Chemistry B* **2011**, *115*, 3434–43.

[65] Jusufi, A.; Sanders, S.; Klein, M. L.; Panagiotopoulos, A. Z. *The Journal of Physical Chemistry B* **2011**, *115*, 990–1001.

[66] Velinova, M.; Sengupta, D.; Tadjer, A. V.; Marrink, S.-J. *Langmuir* **2011**, *27*, 14071–14077.

[67] González-Melchor, M.; Mayoral, E.; Velázquez, M. E.; Alejandre, J. *The Journal of Chemical Physics* **2006**, *125*, 224107.

[68] Lowicki, N. *Fette, Seifen, Anstrichmittel* **1974**, *76*, 136–140.

[69] Sohling, U.; Ruf, F.; Thie, G.; Kuhn, H.; Ralla, K.; Beutel, S.; Kasper, C.; Scheper, T. DE 102011101880 A1. 2011.

[70] King, J. C. *International Journal for Vitamin and Nutrition Research* **2010**, *80*, 300–306.

[71] King, J. *The American Journal of Clinical Nutrition* **2011**, *94*, 679–684.

[72] Schörken, U.; Both, S.; Stuhlmann, D. EP 1925673 A1. 2008.

[73] Kuhn, H.; Ruf, F.; Sohling, U.; Thie, G. EP 2108446 A1. 2009.

[74] Kuhn, H.; Thie, G. WO 2009083069 A2. 2009.

[75] van Meer, G. *Annual Review of Cell Biology* **1989**, *5*, 247–275.

[76] van Meer, G.; Voelker, D. R.; Feigenson, G. W. *Nature Reviews. Molecular Cell Biology* **2008**, *9*, 112–124.

[77] Zachowski, A. *Biochemical Journal* **1993**, *294*, 1–14.

[78] Dowhan, W. *Annual Review of Biochemistry* **1997**, *66*, 199–232.

[79] Shillcock, J. C.; Lipowsky, R. *The Journal of Chemical Physics* **2002**, *117*, 5048.

[80] Marrink, S. J.; Mark, A. E. *Journal of the American Chemical Society* **2003**, *125*, 11144–11145.

[81] Illya, G.; Lipowsky, R.; Shillcock, J. C. *The Journal of Chemical Physics* **2006**, *125*, 114710.

[82] Goetz, R.; Gompper, G.; Lipowsky, R. *Physical Review Letters* **1999**, *82*, 221–224.

[83] Andersen, O. S.; Koeppe, R. E. *Annual Review of Biophysics and Biomolecular Structure* **2007**, *36*, 107–310.

[84] Edidin, M. *Nature Reviews. Molecular Cell Biology* **2003**, *4*, 414–418.

[85] Fujiki, Y.; Hubbard, a. L.; Fowler, S.; Lazarow, P. B. *The Journal of Cell Bology* **1982**, *93*, 97–102.

[86] Schubert, R.; Beyer, K.; Wolburg, H.; Schmidt, K. H. *Biochemistry* **1986**, *25*, 5263–5269.

[87] Tanford, C. *The Journal of Physical Chemistry* **1974**, *2469*, 2469–2479.

[88] Grafmüller, A.; Shillcock, J.; Lipowsky, R. *Biophysical Journal* **2009**, *96*, 2658–2675.

[89] Craik, D. J. *Journal of Peptide Science* **2013**, *19*, 393–407.

[90] Gründemann, C.; Thell, K.; Lengen, K.; Garcia-Käufer, M.; Huang, Y.-H.; Huber, R.; Craik, D. J.; Schabbauer, G.; Gruber, C. W. *PLoS ONE* **2013**, *8*, e68016.

[91] Burman, R.; Gunasekera, S.; Strömstedt, A. a.; Göransson, U. *Journal of Natural Products* **2014**, *77*, 724–736.

[92] Huang, Y.-H.; Colgrave, M. L.; Daly, N. L.; Keleshian, A.; Martinac, B.; Craik, D. J. *The Journal of Biological Chemistry* **2009**, *284*, 20699–707.

[93] Shenkarev, Z. O.; Nadezhdin, K. D.; Sobol, V. A.; Sobol, A. G.; Skjeldal, L.; Arseniev, A. S. *The FEBS Journal* **2006**, *273*, 2658–72.

[94] Daly, N. L.; Rosengren, K. J.; Craik, D. J. *Advanced Drug Delivery Reviews* **2009**, *61*, 918–930.

[95] Wang, C. K.; Wacklin, H. P.; Craik, D. J. *The Journal of Bological Chemistry* **2012**, *287*, 43884–98.

[96] Nawae, W.; Hannongbua, S.; Ruengjitchatchawalya, M. *Scientific Reports* **2014**, *4*, 1–9.

[97] Leach, A. R. *Molecular modelling: principles and applications*; Prentice Hall, 2001; p 768.

[98] Ramachandran, K. I.; Deepa, G.; Namboori, K. *Computational Chemistry and Molecular Modeling: Principles and Applications*; Springer, 2008; p 398.

[99] Del Ben, M.; Schönherr, M.; Hutter, J.; VandeVondele, J. *The Journal of Physical Chemistry Letters* **2013**, *4*, 3753–3759.

[100] Sun, H. *The Journal of Physical Chemistry B* **1998**, *102*, 7338–7364.

[101] Jensen, F. *Introduction to computational chemistry*, 2nd ed.; John Wiley & Sons, Ltd, 2007; p 2007.

[102] Bröer, B. M. *Der Nociceptin-Rezeptor: Theoretische Untersuchungen zu Ligand-Protein-Wechselwirkungen*; Cuvillier Verlag Göttingen, 2004; p 206.

[103] MSI, *Discover, Part 1, General, for Discover 96.0, 4.0.0, 2.9.8*; Molecular Simulations (MSI): San Diego Calif., 1996; p 778.

[104] Polyak, B. T. *USSR Computational Mathematics and Mathematical Physics* **1969**, *9*, 94–112.

[105] Polak, E.; Ribiere, G. *ESAIM: Mathematical Modelling and Numerical ...* **1969**, *3*, 35–43.

[106] Shi, Z.; Guo, J. *Computational and Applied Mathematics* **2008**, *27*, 93–106.

[107] Salomon, R. *IEEE Transactions on Evolutionary Computation* **1998**, *2*, 45–55.

[108] Hünenberger, P. *Advanced Computer Simulation* **2005**, *173*, 105–149.

[109] Andersen, H. C. *The Journal of Chemical Physics* **1980**, *72*, 2384–2393.

[110] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.

[111] Nose, S. *Molecular Physics* **1984**, *52*, 255–268.

[112] Nose, S. *The Journal of Chemical Physics* **1984**, *81*, 511–519.

[113] Verlet, L. *Physical Review* **1967**, *i*, 159–165.

[114] Allen, M.; Tildesley, D. *Computer simulation of liquids*; Clarendon Press, Oxford, 1987; p 400.

[115] Swope, W. C. *The Journal of Chemical Physics* **1982**, *76*, 637.

[116] Duan, Y.; Wu, C.; Chowdhury, S. *Journal of Computational Chemistry* **2003**, *24*, 1999–2012.

[117] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *Journal of Computational Chemistry* **1983**, *4*, 187–217.

[118] MacKerell, A.; Bashford, D. *The Journal of Physical Chemistry B* **1998**, *102*, 3586–3616.

[119] Halgren, T. A. *Journal of the American Chemical Society* **1992**, *114*, 7827–7843.

[120] Groot, R. D.; Warren, P. B. *The Journal of Chemical Physics* **1997**, *107*, 4423.

[121] Marsh, C.; Yeomans, J. *EPL (Europhysics Letters)* **1997**, *37*, 511–516.

[122] Flory, P. J. *The Journal of Chemical Physics* **1942**, *10*, 51.

[123] Huggins, M. *Annals of the New York Academy of Sciences* **1942**, *43*, 1–32.

[124] Fried, J. *Polymer science and technology*; Prentice-Hall, 2003; p 509.

[125] Chen, S.; Doolen, G. *Annual review of fluid mechanics* **1998**, *30*, 329–364.

[126] McNamara, G.; Zanetti, G. *Physical Review Letters* **1988**, *61*, 2332–2335.

[127] Benzi, R.; Succi, S.; Vergassola, M. *Physics Reports* **1992**, *222*, 145–197.

[128] Ermak, D.; Buckholz, H. *Journal of Computational Physics* **1980**, *35*, 169–182.

[129] Ermak, D. L. *Rapport d'activité scientifique du CECAM* **1976**, 66–81.

[130] Frisch, U.; D'Humieres, D.; Hasslbacher, B.; Lallemand, P.; Pomeau, Y.; Rivet, J.-P. *Complex Systems* **1987**, *1*, 649–707.

[131] Groot, R. D.; Madden, T. J.; Tildesley, D. J. *The Journal of Chemical Physics* **1999**, *110*, 9739.

[132] Goga, N.; Rzepiela, A. *Journal of Chemical Theory and Computation* **2012**, *8*, 3637–3649.

[133] Doi, M.; Edwars, S. F. *The theory of polymer dynamics*; Oxford Science Publications: New York, 1988; p 406.

[134] Wikipedia, Biomembran — Wikipedia, Die freie Enzyklopädie. 2014; `\url{http://de.wikipedia.org/wiki/Biomembran}`, [Online: 12. Juni 2014].

[135] Singer, S. J.; Nicolson, G. L. *Science* **1972**, *175*, 720–731.

[136] Nicolson, G. L. *Biochimica et Biophysica Acta* **2014**, *1838*, 1451–1466.

[137] Chernomordik, L. V.; Kozlov, M. M. *Nature Structural & Molecular Biology* **2008**, *15*, 675–683.

[138] Lipowsky, R. *Nature* **1991**, *349*, 475–481.

[139] Wimley, W. C.; Thompson, T. E. *Biochemistry* **1990**, *29*, 1296–1303.

[140] Brown, K. L.; Conboy, J. C. *The Journal of Physical Chemistry B* **2013**, *117*, 15041–15050.

[141] Berg, J. M.; Tymoczko, J. L.; L, S. *Biochemistry*, 6th ed.; Elsevier GmbH: München, 2007; p 1224.

[142] Wikipedia, Aminosäuren — Wikipedia, Die freie Enzyklopädie. 2014; `\url{http://de.wikipedia.org/w/index.php?title=Aminos%C3%A4uren&oldid=126687509}`, [Online: 3. Februar 2014].

[143] Alberts, B.; Johnson, A.; Walter, P.; Lewis, J.; Raff, M.; Roberts, K. *The Yale journal of biology and medicine*, 5th ed.; Taylor & Francis: New York, 2008; p 1392.

[144] Wikipedia, Ramachandran plot — Wikipedia, Die freie Enzyklopädie. 2014; `\url{http://en.wikipedia.org/wiki/Ramachandran_plot}`, [Online: 13. Mai 2014].

[145] Thie, G. Entwicklung und Untersuchung grenzflächenaktiver Zinkrizinoleat-Komplexe zur Adsorption von Schadstoffen aus der Gasphase. Ph.D. thesis, Universität Duisburg-Essen, 2010.

[146] Billault, I.; Mantle, P. G.; Robins, R. J. *Journal of the American Chemical Society* **2004**, *126*, 3250–3256.

[147] Ahmad, F.; Maier, R.; Mukherjee, K. D.; Mangold, H. K. *Lipids* **1987**, *22*, 413–416.

[148] Weiss, J.; Weiss, T. *Handbook of Ion Chromatography*; Wiley-VCH Verlag GmbH & Co. KGaA, 2004; p 931.

[149] Bhattacharyya, L.; Rohrer, J. S. *Applications of Ion Chromatography in the Analysis of Pharmaceutical and Biological Products*; John Wiley & Sons, 2012; p 476.

[150] Lottspeich, F.; Engels, J. W. *Bioanalytik*; Spektrum Akademischer Verlag, 2012; p 1208.

[151] Janson, J.-C.; Ryden, L. *Protein Purification: Principles, High Resolution Methods, and Applications*; John Wiley & Sons, 1998; p 712.

[152] Zhao, C.; Hellman, L. M.; Zhan, X.; Bowman, W. S.; Whiteheart, S. W.; Fried, M. G. *Analytical Biochemistry* **2010**, *399*, 237–245.

[153] Hengen, P. *Trends in Biochemical Sciences* **1995**, *20*, 285–286.

[154] Hochuli, E.; Döbeli, H.; Schacher, A. *Journal of Chromatography A* **1987**, *411*.

[155] Hochuli, E.; Bannwarth, W.; Döbeli, H. *Nature Biotechnology* **1988**, *6*, 1321–1325.

[156] Cormier, A.; Marchand, M.; Ravelli, R. B. G.; Knossow, M.; Gigant, B. *EMBO Reports* **2008**, *9*, 1101–1106.

[157] Porath, J.; Carlsson, J.; Olsson, I.; Belfrage, G. *Nature* **1975**, *258*, 598–599.

[158] Bende, H. *Chemie in unserer Zeit* **1974**, *125*, 17–25.

[159] Barry, D. G.; Daly, N. L.; Clark, R. J.; Sando, L.; Craik, D. J. *Biochemistry* **2003**, *42*, 6688–95.

[160] Herrmann, A.; Svangå rd, E.; Claeson, P.; Gullbo, J.; Bohlin, L.; Göransson, U. *Cellular and Molecular Life Sciences : CMLS* **2006**, *63*, 235–45.

[161] Svangå rd, E.; Burman, R.; Gunasekera, S.; Lövborg, H.; Gullbo, J.; Göransson, U. *Journal of Natural Products* **2007**, *70*, 643–647.

[162] Henriques, S. T.; Huang, Y.-H.; Castanho, M. A. R. B.; Bagatolli, L. A.; Sonza, S.; Tachedjian, G.; Daly, N. L.; Craik, D. J. *The Journal of Biological Chemistry* **2012**, *287*, 33629–43.

[163] Groot, R. D. *The Journal of Chemical Physics* **2003**, *118*, 11265–11277.

[164] Ibergay, C.; Malfreyt, P.; Tildesley, D. J. *Journal of Chemical Theory and Computation* **2009**, *5*, 3245–3259.

[165] Bank, R. P. D. About the PDB Archive and the RCSB PDB. `http://www.rcsb.org/pdb/static.do?p=general_information/about_pdb/index.html`, 2013.

[166] Steinbeck, C.; Han, Y.; Kuhn, S.; Horlacher, O.; Luttmann, E.; Willighagen, E. *Journal of Chemical Information and Computer Sciences* **2003**, *43*, 493–500.

[167] Steinbeck, C.; Hoppe, C.; Kuhn, S.; Floris, M.; Guha, R.; Willighagen, E. L. *Current Pharmaceutical Design* **2006**, *12*, 2111–20.

[168] Callaway, J.; Cummings, M.; Deroski, B.; Esposito, P.; Forman, A.; Langdon, P.; Libeson, M.; McCarthy, J.; Sikora, J.; Xue, D. *Brookhaven National Laboratory* **1996**,

[169] Shaanan, B. *Journal of molecular biology* **1983**, *171*, 31–59.

[170] Cherezov, V.; Rosenbaum, D. M.; Hanson, M. a.; Rasmussen, S. r. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. *Science (New York, N.Y.)* **2007**, *318*, 1258–1265.

[171] Castellani, F.; Rossum, B. V.; Diehl, A. *Nature* **2002**, *420*, 98–102.

[172] Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E. *Science* **1996**, *274*, 1859–1866.

[173] Richardson, J. S.; Richardson, D. C. In *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G. D., Ed.; Plenum Press, 1989; p 798.

[174] Rosen, M. *Surfactants and interfacial phenomena*; John Wiley & Sons, 2004.

[175] Elworthy, P.; MacFarlane, C. *Journal of Pharmacy and Pharmacology* **1965**, *20*, 65–82.

[176] Walton, J.; Tildesley, D.; Rowlinson, J.; Henderson, J. *Molecular Physics* **1983**, *48*, 1357–1368.

[177] Kiselev, M. A.; Zemlyanaya, E. V.; Aswal, V. K.; Neubert, R. H. H. *European Biophysics Journal* **2006**, *35*, 477–93.

[178] Gurtovenko, A. A.; Vattulainen, I. *The Journal of Physical Chemistry. B* **2007**, *111*, 13554–13559.

[179] Gerelli, Y.; Porcar, L.; Lombardi, L.; Fragneto, G. *Langmuir: The ACS Journal of Surfaces and Colloids* **2013**, *29*, 12762–9.

[180] Gao, L.; Shillcock, J.; Lipowsky, R. *The Journal of Chemical Physics* **2007**, *126*, 015101.

[181] Kornberg, R. D.; McConnell, H. M. *Biochemistry* **1971**, *10*, 1111–20.

[182] Hoekstra, D. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1982**, *692*, 171–175.

[183] Kiselev, M. A.; Zemlyanaya, E. V.; Ryabova, N. Y.; Hauss, T.; Almasy, L.; Funari, S. S.; Zbytovska, J.; Lombardo, D. *Applied Physics A* **2013**,

[184] Larsson, K. *Chemistry and Physics of Lipids* **1997**, *88*, 15–20.

[185] Feigenson, G. W. *Annual Review of Biophysics and Biomolecular Structure* **2007**, *36*, 63–77.

[186] Devaux, P. *Annual Review of Biophysics and Biomolecular Structure* **1992**, *21*, 417–439.

[187] Daleke, D. L. *Journal of Lipid Research* **2003**, *44*, 233–242.

[188] Haluska, C. K.; Riske, K. A.; Lehn, J.-M.; Lipowsky, R.; Dimova, R. *PNAS* **2006**, *103*, 15841–15846.

[189] Lindau, M.; Alvarez de Toledo, G. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2003**, *1641*, 167–173.

[190] Noguchi, H.; Takasu, M. *The Journal of Chemical Physics* **2001**, *115*, 9547–9551.

[191] Knecht, V.; Mark, A. E.; Marrink, S.-J. *Journal of the American Chemical Society* **2006**, *128*, 2030–2034.

[192] Knecht, V.; Marrink, S.-J. *Biophysical Journal* **2007**, *92*, 4254–4261.

[193] Wang, X. *Biochemistry* **2012**, *51*, 8353–8362.

[194] Kursula, P. *Acta Crystallographica. Section D, Biological Crystallography* **2014**, *70*, 24–30.

[195] Seaton, B. a.; Head, J. F.; Engelman, D. M.; Richards, F. M. *Biochemistry* **1985**, *24*, 6740–6743.

[196] Chou, J. J.; Li, S.; Klee, C. B.; Bax, A. *Nature Structural Biology* **2001**, *8*, 990–997.

[197] Turkey, J. W. *Exploratory Data Analysis*; Addison Wesley Pub Co Inc, 1977; p 503.

[198] Fermi, G.; Perutz, M. F.; Shaanan, B.; Fourme, R. *Journal of Molecular Biology* **1984**, *175*, 159–174.

[199] Vandegriff, K. D.; McCarthy, M.; Rohlfs, R. J.; Winslow, R. M. *Biophysical Chemistry* **1997**, *69*, 23–30.

[200] Baker, E. N.; Blundell, T. L.; Cutfield, J. F.; Cutfield, S. M.; Dodson, E. J.; Dodson, G. G.; Hodgkin, D. M.; Hubbard, R. E.; Isaacs, N. W.; Reynolds, C. D. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **1988**, *319*, 369–456.

[201] Rosengren, K. J.; Daly, N. L.; Plan, M. R.; Waine, C.; Craik, D. J. *The Journal of Biological Chemistry* **2003**, *278*, 8606–8616.

[202] Kuhn, H.; Müller, F.; Peggau, J.; Zekorn, R. *Journal of Surfactants and Detergents* **2000**, *3*, 335–343.

11 Anhang

11.1 Abkürzungsverzeichnis

Tabelle 11.1: *Liste der verwendeten Abkürzungen.*

Abkürzung	Definition
$\hbar$	Planckkonstante dividiert durch $2 \cdot \pi$
$\mathcal{H}$	Hamilton-Operator
∇	Nabla-Operator
α	Stärke der Adaption
γ	Oberflächenspannung
γ_{CMC}	Oberflächenspannung bei CMC
γ_{DPD}	Reibungskoeffizient
γ_0	Oberflächenspannung der Wasser-Luft Grenzfläche
ΔE_{ij}	Energiedifferenz der paarweisen Wechselwirkung
η	Atomkern-Wellenfunktion
θ	Bindungswinkel
θ	Standardverteilte Zufallsvaribale mit Einheitsvarianz
ξ	Elektrisches Feld
π	Kreiszahl
Π_{CMC}	Effektivität der Reduktion der Oberflächenspannung

Abkürzung	Definition
σ	Amplitude der stochastischen Kraft
ϕ	Elektronen-Wellenfunktion
ϕ	Torsionswinkel
χ	*out-of-plane* Winkel
χ_{ij}	Flory-Huggins-Parameter
Ψ	Wellenfunktion eines Teilchens
ω^D	Variable des Reibungskoeffizienten
ω^R	Variation der zufälligen Kraft mit der Distanz
a	Beschleunigung
a_{ij}	DPD-Replusionsparameter
a_k	Schrittweite
AMBER	*Assisted Model Building with Energy Refinement*
API	*Application Programming Interface*
b	Bindung
$c_{correct}$	Korrekturfaktor
$c_{damping}$	Dämpfungsfaktor
$c_{effective}$	Korrekturfaktor der effektiven Ladung
BCIs (δ_{ij})	*Bond-Charge Increments*
CCK	*Cyclic Cystine Knot*
CDK	*Chemistry Development Kit*
CG	*Conjugate-Gradient*
CHARMM	*Chemistry at HARvard Macromolecular Mechanics*
CMC	Kritische Mizellenkonzentration
COMPASS	*Condensed-Phase Optimized Molecular Potentials for Atomistic Simulation Studies*

Abkürzung	Definition
DMPC	1,2-Dimyristoyl-sn-glycero-3-PC
DOPC	1,2-Dioleoyl-sn-glycero-3-PC
DPD	Dissipative Partikel Dynamik
DPPC	1,2-Dipalmitoyl-sn-glycero-3-PC
E	Energie
E_e	Bindungswinkelenergie
E_{el}	Elektrostatische Energie
E_k	*Cross-term* Energie
E_{kin}	Kinetische Energie
E_{KF}	Gesamtenergie eines Kraftfeldes
E_l	Bindungslängenenergie
E_t	Torsionswinkelenergie
E_{vdw}	Van-der-Waals-Energie
ER	Endoplasmatisches Retikulum
F	Kraft
F^C	Konservative Kraft
F^D	Dissipative Kraft
F^E	Elektrostatische Kraft
$\underline{F}^E_{max}$	Maximale elektrostatische Kraft
F^R	Zufällige Kraft
GUI	*Graphical User Interface*
HF	Hartree-Fock-Näherung
HP	*Hydrophobic Patch*
i	Imaginäre Einheit

Abkürzung	Definition
IMAC	*Immobilized Metal Ion Affinity Chromatography*
k_b	Boltzmann-Konstante
k_{BB}	Kraftkonstante der Stabilisierungspotentiale
KLP	Kombination von kurz- und langreichweitigen Potentialen
KP	Kurzreichweitige Potentiale
$LD50_{oral}$	Mittlere orale letale Dosis
LJ	Lennard-Jones-Potential
LP	Langreichweitige Potentiale
m	Masse
MC	Monte-Carlo
MD	Molekulardynamik
MFD	Molekulare Fragmentdynamik
MFT	Molekularfeldtheorie
MM	Molekularmechanik
N	Partikelanzahl
$N_a a$	Anzahl der Aminosäuren
N_p	Anzahl der Stabilisierungspotentiale
NVE	Mikrokanonisches Ensemble
NVT	Kanonisches Ensemble
p	Impuls
PC	Phosphatidylcholin
PCFF	Polymer Consistent Force Field
PDB	Protein Data Bank
PE	Phosphatidylethanolamin
PI	Phosphatidylinositol

Abkürzung	Definition
POPC	2-Oleoyl-1-palmitoyl-sn-glycero-PC
PS	Phosphatidylserin
q	Partialladung
q^*	Effektive Ladung
QM	Quantenmechanik
QSAR	*Quantitative Structure-Activity Relationship*
r	Distanz- oder Ortsvektor
R	Universelle Gaskonstante
R^n	n-dimensionaler euklidischer Raum
r_{cutoff}	*Cutoff* Radius
$r_{damping}$	Dämpfungsabstand
R_g	Gyrationsradius
$R_{g,fragment}$	Fragmentbasierter Gyrationsradius
SD	*Steepest-Descent*
t	Zeit
$T(t)$	Aktuelle Temperatur
T_{md}	Vorgegebene Temperatur
v	Geschwindigkeit
V	Volumen
VdW	Van-der-Waals Wechselwirkung
x_0	Startpunkt
Z	Koordinationszahl
$Zn(Ri)_2$	Zinkricinoleat

11.2 Kodierungsschema der Aminosäuren

Tabelle 11.2: *Unterstützte Aminosäuren.*

Aminosäure	Dreibuchstabencode	Einbuchstabencode
Alanin	Ala	A
Arginin	Arg	R
Asparagin	Asn	N
Asparaginsäure	Asp	D
Cystein	Cys	C
Glutamin	Gln	Q
Glutaminsäure	Glu	E
Glycin	Gly	G
Histidin	His	H
Isoleucin	Ile	I
Leucin	Leu	L
Lysin	Lys	K
Methionin	Met	M
Phenylalanin	Phe	F
Prolin	Pro	P
Serin	Ser	S
Threonin	Thr	T
Tryptophan	Trp	W
Tyrosin	Tyr	Y
Valin	Val	V

Tabelle 11.3: *Kodierungschema der Aminosäuren.*

Aminosäure	*f*Smiles	Ladungskodierung
Alanin	MeAcNH2(Me)	TC&2.3?HAc:Ac-;TN&9.9?MeNH3+:MeNH2;
Arginin	MeAcNH2(Pr-Guanidine)	TC&2.81?HAc:Ac-;TN&9.09?MeNH3+:MeNH2; SCF&12.48?GuanidineH+:Guanidine;SCB&12.48?GuanidineP+:Guanidine;
Asparagin	MeAcNH2(AcNH2)	TC&2.02?HAc:Ac-;TN&8.8?MeNH3+:MeNH2;
Asparaginsäure	MeAcNH2(HAc)	TC&1.88?HAc:Ac-;TN&9.6?MeNH3+:MeNH2;SCF&3.68?HAc:Ac-;SCB&3.85?Ac:Ac-;
Cystein	MeAcNH2(MeSH)	TC&1.71?HAc:Ac-;TN&10.78?MeNH3+:MeNH2; SCF&8.33?MeSH:MeS-;SCB&8.95?MeSH:MeS-;SS&MeSH?MeSHSS;
Glutaminsäure	MeAcNH2(Me-HAc)	TC&2.19?HAc:Ac-;TN&9.67?MeNH3+:MeNH2; SCF&4.25?HAc:Ac-;SCB&4.35?HAc:Ac-;
Glutamin	MeAcNH2(Me-AcNH2)	TC&2.17?HAc:Ac-;TN&9.13?MeNH3+:MeNH2;
Glycin	MeAcNH2	TC&2.21?HAc:Ac-;TN&9.15?MeNH3+:MeNH2
Histidin	MeAcNH2(Me-Imidazole)	TC&1.78?HAc:Ac-;TN&8.97?MeNH3+:MeNH2;

Aminosäure	*f*Smiles	Ladungskodierung
		SCF&6.0?ImidazoleH+:Imidazole;SCB&6.9?ImidazoleH+:Imidazole;
Isoleucin	MeAcNH2(Et-Et)	TC&2.32?HAc:Ac-;TN&9.76?MeNH3+:MeNH2;
Leucin	MeAcNH2(Me-Pr)	TC&2.4?HAc:Ac-;TN&9.6?MeNH3+:MeNH2;
Lysin	MeAcNH2(Pr-MeNH2)	TC&2.2?HAc:Ac-;TN&8.9?MeNH3+:MeNH2; SCF&10.53?MeNH3+:MeNH2;SCB&9.4?MeNH3+:MeNH2
Methionin	MeAcNH2(Et-MeSH)	TC&2.28?HAc:Ac-;TN&9.21?MeNH3+:MeNH2;
Phenylalanin	MeAcNH2(Me-Bz)	TC&2.58?HAc:Ac-;TN&9.24?MeNH3+:MeNH2;
Prolin	ProRing	TC&1.99?HAc:Ac-;TN&10.6?ProRingH+:ProRing;
Serin	MeAcNH2(MeOH)	TC&2.21?HAc:Ac-;TN&9.15?MeNH3+:MeNH2;
Threonin	MeAcNH2(PrOH)	TC&2.1?HAc:Ac-;TN&9.12?MeNH3+:MeNH2;
Tryptophan	MeAcNH2(Me-Pyrrole-Bz)	TC&2.15?HAc:Ac-;TN&9.12?MeNH3+:MeNH2;
Tyrosin	MeAcNH2(Me-Phenol)	TC&2.2?HAc:Ac-;TN&9.11?MeNH3+:MeNH2; SCF&10.07?Phenol:PhenolH-;SCB&9.9?Phenol:PhenolH-;

Aminosäure	*f*Smiles	Ladungskodierung
Valin	MeAcNH2(Pr)	TC&2.3?HAc:Ac-;TN&9.6?MeNH3+:MeNH2;

11.3 Regeln der *f*SMILES-Notation

Folgende Regeln sind für den Aufbau eines *f*SMILES definiert:

1. Fragmentnamen dürfen eine maximale Länge von 10 Buchstaben (a-z, A-Z, 0-9, erstes Zeichen keine Zahl) besitzen und mit einer optionalen vorangestellten Häufigkeitszahl versehen sein.

2. Das Zeichen für die chemische Bindung der Hauptkette ist '-'.

3. Runde Klammern '(' und ')' bestimmen Verzweigungen. Diese können in Abhängigkeit von der Hierarchieebene verschachtelt werden.

4. Eckige Klammern '[' und ']' mit einer eingeschlossenen Nummer im Anschluss eines Fragmentes zeigen einen Ringschluss.

5. Geschweifte Klammern '{' und '}' definieren ein Monomer. Monomere bestehen aus mindestens einem Fragment mit einem [HEAD]- und [TAIL]-Attribut: Strukturelement die dem Monomer vorangehen, werden mit dem [HEAD]-Fragment, nachfolgende Elemente mit dem [TAIL]-Fragment verknüpft. Verschachtlungen sind nicht erlaubt.

6. Monomerlabel beginnen mit einem '#'-Zeichen, gefolgt von einer Zeichenkette (erstes Zeichen keine Zahl).

7. Monomerlabels können mit einer voranstehenden Häufigkeitszahl versehen werden.

8. Ein Fragment einer Molekülstruktur (nicht-Monomer) kann optional ein [START]- oder ein [END]-Attribut für Ausrichtungszwecke beinhalten. Nur ein [START]/[END]-Paar ist pro Struktur erlaubt.

9. Ein Molekül kann aus mehreren distinkten Teilen bestehen. Jeder Teil muss von spitzen Klammern '<' und '>' eingefasst werden. Verschachtlungen sind nicht erlaubt.

11.4 Lebenslauf

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Publikationen

Truszkowski, A., Jayaseelan, K. V., Neumann, S., Willighagen, E. L., Zielesny, A., & Steinbeck, C. (2011). *New developments on the cheminformatics open workflow environment CDK-Taverna.* Journal of Cheminformatics, 3, 54.

Jayaseelan, K. V., Moreno, P., Truszkowski, A., Ertl, P., & Steinbeck, C. (2012). *Natural product-likeness score revisited: an open-source, open-data implementation.* BMC Bioinformatics, 13, 106.

Für die vorliegende Arbeit:

Truszkowski, A., Epple, M., Fiethen, A., Zielesny, A., & Kuhn, H. (2013). *Molecular fragment dynamics study on the water-air interface behavior of non-ionic polyoxyethylene alkyl ether surfactants.* Journal of Colloid and Interface Science, 410, 140–145.

Truszkowski, A., Daniel, M., Kuhn, H., Neumann, S., Steinbeck, C., Zielesny, A. & Epple, M. (2014). *A Molecular Fragment Cheminformatics Roadmap for Mesoscopic Simulation.* Journal of Cheminformatics, 6, 45.

Truszkowski, A., van den Broek, K., Kuhn, H., Zielesny, A. & Epple, M. (2014). *Mesoscopic simulation of biological membranes with Molecular Fragment Dynamics.*, Journal of Chemical Information and Modeling, submitted.

Truszkowski, A., van den Broek, K., Kuhn, H., Zielesny, A. & Epple, M. (2014). *Mesoscopic simulation of peptides and proteins with Molecular Fragment Dynamics.*, Journal of Chemical Information and Modeling, submitted.

11.5 Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

"Molekulare Simulation von Proteinen, Peptiden und Biomembranen mit Molekularer Fragmentdynamik (MFD)"

selbst verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe, und dass die Arbeit in dieser oder ähnlicher Form noch bei keiner anderen Universität eingereicht wurde.

Essen, 13. November 2014

11.6 Danksagung

"What we do in life echoes in eternity."

Gladiator (2000 Film)

An dieser Stelle möchte ich mich bei allen bedanken, die diese Arbeit ermöglicht und unterstützt haben.

Mein besonderer Dank gilt Prof. Dr. Matthias Epple für die sehr gute Betreuung und stete Diskussionsbereitschaft, die diese Arbeit möglich gemacht haben, sowie die Freiräume, die mir gelassen wurden.

Des weiteren möchte ich mich bei Prof. Dr. Achim Zielesny für die Übernahme des Zweitgutachtens sowie für die umfassende Beratung im chemo- und bioinformatischen Bereich bedanken. Ebenso möchte ich mich für das private Engagement bedanken, das er hinsichtlich meiner Ausbildung und Förderung gezeigt hat.

Zudem gilt mein Dank den Mitarbeitern von Cam-D Technologies GmbH für die Unterstützung bei Fragen im Bereich des Molecular Modelling und der molekularen Computersimulation.

Gleichermaßen möchte ich mich bei der Gesellschaft für naturwissenschaftliche Informatik mbH für diese großartige Zusammenarbeit und Unterstützung in software-technischen Fragen bedanken.

Darüber hinaus gehört mein besonders herzlicher Dank meinen Eltern Johann und Doris Truszkowski sowie meiner Lebensgefährtin Josefine Graw, die mein Studium und meine Dissertation ermöglicht sowie mich währenddessen voll und ganz unterstützt haben.